V.I. ARNOLD

LES MÉTHODES MATHÉMATIQUES DE LA MÉCANIQUE CLASSIQUE

V. ARNOLD

LES MÉTHODES MATHÉMATIQUES DE LA MÉCANIQUE CLASSIQUE

ÉDITIONS MIR · MOSCOU

TABLE DES MATIÈRES

TROISIÈME PARTIE

MÉCANIQUE DE HAMILTON

AVANT-PROPOS

La mécanique classique utilise un arsenal très riche de méthodes mathématiques et de notions: équations différentielles, flots, applications et variétés différentiables, groupes et algèbres de Lie, géométrie symplectique et théorie ergodique. Nombre de théories mathématiques modernes doivent leur existence à des problèmes de mécanique et ce n'est que par la suite qu'elles ont acquis cette forme axiomatique et abstraite qui complique tant leur étude.

L'appareil mathématique de la mécanique classique est exposé dans son intégrité, si bien que l'on ne supposera au lecteur que des connaissances préliminaires ne dépassant pas le cadre des traditionnels cours d'analyse (dérivée, intégrale, équations différentielles ordinaires), de géométrie (espace vectoriel, vecteurs) et d'algèbre linéaire (opérateurs linéaires, formes quadratiques).

Cet appareil nous sera grandement utile pour développer les problèmes fondamentaux de la dynamique d'un système et notamment la théorie des oscillations, la théorie du mouvement d'un solide et le formalisme hamiltonien. L'auteur s'est partout efforcé à mettre en exergue l'aspect géométrique, qualitatif des phénomènes. Vu sous cet angle, cet ouvrage est plus proche des cours de mécanique théorique pour les physiciens théoriciens que des cours traditionnels de mécanique théorique pour les mathématiciens.

Une partie importante est consacrée aux principes variationnels et à la dynamique analytique, dynamique analytique à propos de laquelle dans ses *Leçons sur l'évolution des mathématiques au* XIXe *siècle* F. Klein écrivait: « le physicien ne peut tirer que peu de choses de ces théories pour ses problèmes, l'ingénieur — rien ». L'essor de la science est venu vigoureusement infirmer cette remarque. Le formalisme hamiltonien est à la base de la mécanique quantique et constitue aujourd'hui l'un des outils mathématiques le plus souvent mis à contribution en physique. Dès que l'on eut pleinement saisi la portée de la structure symplectique et du principe de Huygens pour tous les problèmes possibles d'optimisation, les équations de Hamilton furent constamment sollicitées dans les

calculs techniques. Par ailleurs le développement actuel de la mécanique céleste, né des besoins des recherches spatiales, a ravivé l'intérêt pour les méthodes et problèmes de la dynamique analytique.

Les Appendices sont consacrés aux imbrications riches et variées de la mécanique classique et des autres domaines des mathématiques.

A titre d'applications de l'appareil de la mécanique on étudie ici les fondements de la géométrie riemannienne, la dynamique du fluide parfait, la théorie de Kolmogorov des perturbations des mouvements quasi périodiques, les asymptotiques à ondes courtes pour les équations de physique mathématique et la classification des caustiques en optique géométrique.

Ces Appendices s'adressent à l'élève désireux d'approfondir ses connaissances. Certains d'entre eux peuvent même faire l'objet de cours spéciaux (nous avons en vue les méthodes asymptotiques de la théorie des oscillations non linéaires ou encore les asymptotiques quasi classiques). Si l'auteur s'est fixé pour principe dans les chapitres fondamentaux de cet ouvrage de faire dans le détail toutes les démonstrations en évitant au possible de renvoyer à d'autres sources, dans les Appendices en revanche il n'a donné que des résultats bruts dont la preuve figure dans les ouvrages de référence.

La base de ce livre est le cours de mécanique classique étalé sur un an et demi que l'auteur a fait aux élèves de 3-ème et 4-ème années de la section de mécanique mathématique de l'Université de Moscou en 1966-1968.

L'auteur tient à exprimer sa reconnaissance au regretté I. Pétrovski qui a insisté pour que ce cours soit lu et édité ; à L. Bounimovitch, L. Weingortine et V. Novikov qui ont mis leurs notes à sa disposition et surtout à N. Kolesnikov qui s'est chargé de la publication des cours polycopiés (1968) ; à tous ses auditeurs et collègues qui lui ont communiqué leurs remarques sur le texte polycopié, remarques dont la plupart ont été utilisées dans la préparation de cette édition ; et enfin à M. Léontovitch qui lui a proposé d'interprêter les liaisons à l'aide du passage à la limite ainsi qu'à I. Vorovitch et V. Youdovitch pour leur attentive lecture du manuscrit.

V. Arnold

PREMIÈRE PARTIE

MÉCANIQUE NEWTONIENNE

La mécanique newtonienne a pour objet l'étude du mouvement d'un système de points matériels dans un espace euclidien de dimension 3. Sur cet espace opère un groupe de mouvements de dimension 6. Les principales notions et théorèmes généraux de la mécanique newtonienne (même formulés en termes de coordonnées cartésiennes) sont invariants par ce groupe *).

Un système mécanique potentiel newtonien est défini par les masses de ses points et par son énergie potentielle. Les lois de conservation correspondent aux mouvements qui laissent invariante l'énergie potentielle.

Les équations de Newton permettent d'étudier complètement un grand nombre de problèmes importants de mécanique, tel le mouvement dans un champ central.

*) Voire par un groupe plus vaste de transformations galiléennes de l'espace-temps.

CHAPITRE PREMIER

FAITS D'EXPÉRIENCES

Dans ce chapitre sont décrits les principaux faits d'expériences qui constituent la base de la mécanique: le principe de relativité de Galilée et l'équation différentielle de Newton. Sont également étudiées les restrictions imposées aux équations de mouvement par le principe de relativité. Des exemples simples sont traités.

§ 1. Principes de relativité et de déterminisme

Dans ce paragraphe on introduit et discute la notion de système de coordonnées inertial. La formulation mathématique exacte des assertions énoncées dans ce paragraphe sera donnée dans le suivant.

La mécanique classique repose sur de nombreux faits d'expériences *). Citons quelques-uns d'entre eux.

A. L'espace et le temps. Notre espace est l'espace euclidien ordinaire de dimension 3, le temps est de dimension 1.

B. Le principe de relativité de Galilée. Il existe des systèmes de coordonnées (dits inertiaux) jouissant des deux propriétés suivantes:

1. Toutes les lois de la nature sont les mêmes à toutes les dates dans tous les systèmes de coordonnées inertiaux.
2. Tous les systèmes de coordonnées se déplaçant en mouvement rectiligne uniforme par rapport à un système inertial sont inertiaux.

En d'autres termes, si un système de coordonnées lié à la Terre est inertial, alors un observateur, qui se trouverait dans un train en mouvement rectiligne uniforme par rapport à la Terre, ne mettrait pas en évidence ce mouvement par les seules expériences effectuées à l'intérieur d'un wagon.

---

*) Tous ces « faits d'expériences » ne sont vrais qu'approximativement: ils ne résistent pas à des expériences plus précises. Pour éviter les expressions encombrantes nous ne le spécifierons pas dans la suite et nous traiterons nos modèles mathématiques comme s'ils décrivaient exactement des phénomènes physiques.

En fait un système de coordonnées lié à la Terre n'est inertial qu'approximativement. Les systèmes de coordonnées liés au Soleil, aux étoiles, etc. sont inertiaux avec une meilleure précision.

C. Principe de déterminisme de Newton. L'état initial d'un système mécanique (l'ensemble des positions et des vitesses de ses points à une date quelconque) définit de façon unique le futur de son mouvement.

Ce fait n'est guère pour nous étonner, car nous le connaissons de bonne heure. On peut s'imaginer un monde dans lequel la définition du futur d'un système impliquerait la connaissance de l'accélération à la date initiale. L'expérience montre que le nôtre est différent.

§ 2. Le groupe de Galilée et les équations de Newton

Dans ce paragraphe nous définirons et étudierons le groupe de transformations de Galilée dans l'espace-temps. Puis nous verrons l'équation de Newton et les restrictions les plus simples imposées à son second membre par les propriétés d'invariance par les transformations de Galilée *).

A. Notations. **R** désignera l'ensemble des réels, $\mathbf{R}^n$ un espace vectoriel réel linéaire de dimension n.

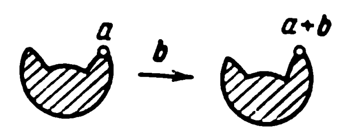


Fig. 1. Translation.

Un *espace affine A^n de dimension n* se distinguera de $\mathbf{R}^n$ par le fait que dans A^n l'« origine des coordonnées n'est pas fixée ». Tout groupe de $\mathbf{R}^n$ opère sur A^n comme un *groupe de translations* (fig. 1):

$$a \to a + \boldsymbol{b} \quad (a \in A^n, \quad \boldsymbol{b} \in \mathbf{R}^n, \quad a + \boldsymbol{b} \in A^n).$$

Donc la somme de deux points de A^n n'est pas définie, contrairement à la différence qui est un vecteur de $\mathbf{R}^n$.

Une *structure euclidienne* dans l'espace vectoriel $\mathbf{R}^n$ est une forme bilinéaire symétrique définie positive appelée *produit scalaire.* Le produit scalaire permet de définir la *distance*

$$\rho(x, y) = \| x - y \| = \sqrt{(x - y, x - y)}$$

de deux points de l'espace *affine* correspondant A^n.

Un espace affine muni d'une telle distance s'appelle *espace euclidien* et se note E^n.

B. Structure de Galilée. La structure spatio-temporelle de Galilée est composée de trois éléments suivants:

*) La lecture de ce paragraphe peut être omise par ceux qui ne ressentent pas le besoin des formulations mathématiques des assertions du § 1.

1. L'*Univers* ou l'espace-temps qui est un espace affine *) de dimension 4 que nous désignerons par A^4. Les points de A^4 sont appelés *points d'Univers* ou *événements*. Les translations de A^4 engendrent un espace vectoriel $\mathbf{R}^4$.

2. Le *temps* qui est une application linéaire $t: \mathbf{R}^4 \to \mathbf{R}$ de l'espace vectoriel $\mathbf{R}^4$ des translations de l'Univers sur l'« axe réel du temps ».

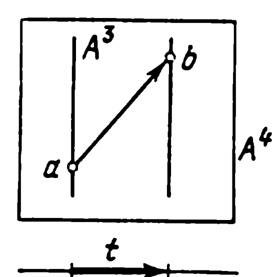


Fig. 2. Intervalle de temps t.

On appelle *intervalle de temps* de l'événement $a \in A^4$ à l'événement $b \in A^4$ le nombre $t(b - a)$ (fig. 2). Si $t(b - a) = 0$, les événements a et b sont *simultanés*.

Tout ensemble d'événements simultanés forme dans A^4 un sous-espace affine A^3 de dimension 3 appelé *espace des événements simultanés*.

Le noyau de l'application t est constitué des translations de A^4 qui transportent un événement (et partant un événement quelconque) dans un événement simultané. Ce noyau est un sous-espace vectoriel $\mathbf{R}^3$ de l'espace vectoriel $\mathbf{R}^4$.

La structure galiléenne comprend encore un élément.

3. La *distance de deux événements simultanés*

$$\rho(a, b) = \| a - b \| = \sqrt{(a - b, a - b)}, \quad a, b \in A^3,$$

définie par le produit scalaire sur $\mathbf{R}^3$. Cette distance transforme tout espace d'événements simultanés en un espace euclidien E^3.

L'espace A^4 muni d'une structure spatio-temporelle de Galilée est appelé *espace de Galilée*.

On pourrait certes parler de deux événements ayant lieu simultanément dans des endroits différents, cependant l'expression « deux événements $a, b \in A^4$ se déroulant à des dates différentes *dans un même endroit d'un espace tridimensionnel* » n'a pas de sens tant qu'on n'a pas fait choix d'un référentiel.

On appelle *groupe de Galilée* le groupe de toutes les transformations conservant la structure de l'espace galiléen dans lequel elles évoluent. Les éléments de ce groupe sont appelés *transformations galiléennes*. Donc les transformations galiléennes sont des transformations affines de A^4 conservant les intervalles de temps et les distances des événements simultanés.

*) Dans l'Antiquité l'Univers était muni d'une structure non pas affine, mais linéaire (un système géocentrique + création du monde).

E x e m p l e. Soit le produit cartésien *) $\mathbf{R} \times \mathbf{R}^3$ de l'axe t par l'espace vectoriel $\mathbf{R}^3$ de structure euclidienne fixée. Un tel espace possède une structure galiléenne naturelle. On l'appellera *espace arithmétique galiléen.*

Citons trois exemples de transformations galiléennes de cet espace. D'abord un *mouvement uniforme de vitesse* v

$$g_1(t, x) = (t, x + vt), \quad \forall\, t \in \mathbf{R}, \quad x \in \mathbf{R}^3.$$

Ensuite une *translation de l'origine des coordonnées*

$$g_2(t, x) = (t + s, x + s), \quad \forall\, t \in \mathbf{R},\ x \in \mathbf{R}^3.$$

Enfin une *rotation des axes de coordonnées*

$$g_3(t, x) = (t, Gx), \quad \forall\, t \in \mathbf{R},\ x \in \mathbf{R}^3,$$

où $G\colon \mathbf{R}^3 \to \mathbf{R}^3$ est une transformation orthogonale.

E x e r c i c e. Montrer que *toute transformation galiléenne de l'espace* $\mathbf{R} \times \mathbf{R}^3$ *peut être représentée de façon unique par le produit d'une rotation, d'une translation et d'un mouvement uniforme* ($g = g_1 \cdot g_2 \cdot g_3$) (si bien que la dimension du groupe de Galilée est égale à $3 + 4 + 3 = 10$).

E x e r c i c e. Montrer que *tous les espaces galiléens sont isomorphes l'un à l'autre* **) *et en particulier sont isomorphes à l'espace arithmétique* $\mathbf{R} \times \mathbf{R}^3$.

Soit un ensemble M. La bijection $\varphi_1\colon M \to \mathbf{R} \times \mathbf{R}^3$ s'appelle *système de coordonnées galiléen sur l'ensemble* M. Un système de coordonnées φ_2 *se déplace uniformément* par rapport au système φ_1 si $\varphi_1 \cdot \varphi_2^{-1}\colon \mathbf{R} \times \mathbf{R}^3 \to \mathbf{R} \times \mathbf{R}^3$ est une transformation galiléenne. Les systèmes de coordonnées galiléens φ_1 et φ_2 définissent une même structure galiléenne sur M.

C. Mouvement, vitesse, accélération. Un mouvement dans $\mathbf{R}^N$ est une application différentiable $x\colon I \to \mathbf{R}^N$ d'un intervalle I de l'axe réel dans $\mathbf{R}^N$.

On appelle *vecteur vitesse* au point $t_0 \in I$ la dérivée

$$\dot{x}(t_0) = \frac{dx}{dt}\Big|_{t=t_0} = \lim_{h \to 0} \frac{x(t_0 + h) - x(t_0)}{h} \in \mathbf{R}^N.$$

On appelle *vecteur accélération* au point t_0 la dérivée seconde

$$\ddot{x}(t_0) = \frac{d^2x}{dt^2}\Big|_{t=t_0}.$$

---

*) On rappelle que le produit cartésien de deux ensembles A, B est l'ensemble des couples ordonnés (a, b), où $a \in A$ et $b \in B$. La structure du produit cartésien est celle de ses termes.

**) i. e. il existe une transformation bijective conservant la structure galiléenne.

Nous admettrons que les fonctions sur lesquelles nous opérerons sont différentiables autant de fois qu'il le faut. Dans la suite, sauf mention du contraire, par applications, fonctions, etc., nous entendrons des applications, des fonctions, etc., différentiables. L'image de l'application $x: I \to \mathbf{R}^N$ est appelée *trajectoire* ou *courbe* de $\mathbf{R}^N$.

E x e r c i c e. Est-ce que la trajectoire d'un mouvement différentiable sur le plan peut affecter la forme représentée sur la figure 3? Le vecteur accélération peut-il prendre la valeur indiquée?

Réponse. Oui. Non.

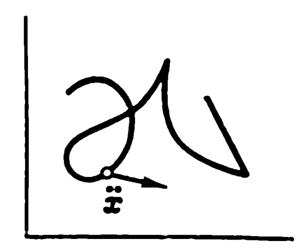


Fig. 3. Trajectoire du mouvement d'un point.

Fig. 4. Lignes d'Univers.

Définissons maintenant un *système mécanique de n points en mouvement dans un espace euclidien de dimension* 3.

Soit $x: \mathbf{R} \to \mathbf{R}^3$ un mouvement dans $\mathbf{R}^3$. Le graphe *) de cette application est une courbe dans $\mathbf{R} \times \mathbf{R}^3$.

On appelle *ligne d'Univers* (fig. 4) une courbe qui, dans un système (et partant un système quelconque) de coordonnées galiléen est le graphe d'un mouvement.

Le mouvement d'un système de n points est défini par n lignes d'Univers dans l'espace galiléen. Dans un système de coordonnées galiléen elles sont décrites par n applications $x_i: \mathbf{R} \to \mathbf{R}^3$, $i = 1, \ldots, n$.

On appelle *espace de configuration* d'un système de n points le produit cartésien de n exemplaires $\mathbf{R}^3$. Nos n applications $x_i: \mathbf{R} \to \mathbf{R}^3$ définissent une application

$$x: \mathbf{R} \to \mathbf{R}^N, \quad N = 3n,$$

de la droite temporelle dans l'espace de configuration. Cette application s'appelle *mouvement d'un système de n points dans un système de coordonnées galiléen* de $\mathbf{R} \times \mathbf{R}^3$.

D. Equation de Newton. En vertu du principe de déterminisme (ou de causalité) de Newton (§ 1, C), le mouvement tout entier d'un

*) On appelle graphe de l'application $f: A \to B$ un sous-ensemble du produit cartésien $A \times B$ constitué de tous les couples de la forme $(a, f(a))$, $a \in A$.

système est univoquement défini par sa position initiale ($x(t_0) \in \mathbf{R}^N$) et ses vitesses initiales ($\dot{x}(t_0) \in \mathbf{R}^N$).

En particulier, la *position et la vitesse initiales définissent l'accélération.* En d'autres termes, il existe une fonction $F: \mathbf{R}^N \times \mathbf{R}^N \times \times \mathbf{R} \to \mathbf{R}^N$ telle que

$$\ddot{x} = F(x, \dot{x}, t). \tag{1}$$

L'équation (1) a été posée à la base de la mécanique par Newton. On l'appelle *équation de Newton.*

En vertu du théorème d'existence et d'unicité de la théorie des équations différentielles ordinaires, la fonction F et les conditions initiales $x(t_0)$, $\dot{x}(t_0)$ définissent univoquement le mouvement *).

La forme de la fonction F est définie empiriquement pour chaque système mécanique concret. Au point de vue mathématique, la forme de F *définit* en soi le système.

E. Restrictions imposées par le principe de relativité. Le principe de relativité de Galilée affirme que l'espace-temps physique est muni d'une structure galiléenne privilégiée (la « classe des systèmes de coordonnées inertiaux ») jouissant de la propriété suivante.

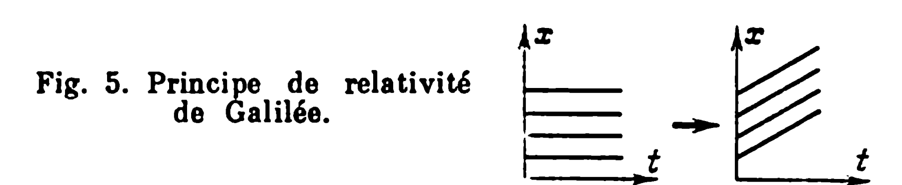


Fig. 5. Principe de relativité de Galilée.

Si les lignes d'Univers de tous les points d'un système mécanique quelconque **) *subissent la même transformation galiléenne, on obtient les lignes d'Univers du même système* (*avec de nouvelles conditions initiales*) (fig. 5).

*) Sous réserve que certaines conditions de différentiabilité soient remplies. Le mouvement n'est défini par l'équation (1) en général que sur un certain *intervalle* de temps. Nous supposerons par souci de simplicité que cet intervalle couvre l'axe temporel tout entier, ce qui est le cas dans la plupart des problèmes de mécanique.

**) Le principe de relativité ne concerne que les systèmes physiques fermés (les systèmes mécaniques notamment), i. e. on devra inclure dans ce système tous les corps qui, par leur interaction, participent à l'événement étudié. En toute rigueur, il faudrait inclure dans le système tous les corps de l'Univers. Cependant l'expérience montre que souvent on peut négliger l'action de beaucoup d'entre eux : par exemple, lorsqu'on étudie le mouvement des planètes autour du Soleil, on peut négliger l'attraction des autres étoiles, etc.

D'un autre côté, lorsqu'on étudie le mouvement des corps au voisinage de la Terre, le système n'est pas fermé s'il ne comprend pas la Terre ; s'agissant du mouvement d'un avion, le système n'est pas fermé si on ne lui ajoute pas l'air environnant l'avion, etc. Dans la suite, par système mécanique on entendra, la plupart des cas, un système fermé : s'il est question de systèmes non fermés, on le spécifiera (cf. par exemple § 3).

2—01408

Ce qui impose au second membre de l'équation de Newton écrite dans un système de coordonnées inertial toute une série de conditions : en particulier, l'équation (1) doit être invariante par le groupe de transformations galiléennes.

Exemple 1. La translation dans le temps est une transformation galiléenne. L'invariance par les translations dans le temps signifie que les « lois de la nature sont constantes », i.e. si $x = \varphi(t)$ est solution de l'équation (1), alors $x = \varphi(t + s)$ le sera également quel que soit $s \in \mathbf{R}$.

Il en résulte que *le second membre de l'équation* (1) *ne peut pas dépendre du temps dans un système de coordonnées inertial*, i.e.

$$\ddot{x} = \Phi(x, \dot{x}).$$

Remarque. Les équations différentielles dont les seconds membres dépendent du temps se rencontrent dans le cas suivant.

Supposons que nous étudions la partie I d'un système mécanique I + II. On peut remplacer parfois l'influence de la partie II sur la partie I par une variation dans le temps des paramètres du système d'équations décrivant le mouvement de la partie I.

Exemple : on peut négliger l'influence de la Lune sur la Terre dans la plupart des phénomènes terrestres. Cependant lorsqu'on étudie les marées, il faut tenir compte de cette influence et remplacer l'attraction de la Lune par une variation périodique de la force de gravité terrestre.

Les équations à coefficients variables peuvent aussi apparaître à la suite d'opérations formelles effectuées pendant la résolution des problèmes.

Exemple 2. Parmi les transformations galiléennes figurent les translations dans un espace de dimension 3. L'invariance par ces translations signifie que cet espace est *homogène* ou encore qu'il « possède les mêmes propriétés en tous ses points ». C'est-à-dire, si $x_i = \varphi_i(t)$ $(i = 1, \ldots, n)$ est le mouvement d'un système de n points vérifiant (1), alors le mouvement $\varphi_i(t) + r$ $(i = 1, \ldots, n)$ vérifie également l'équation (1) quel que soit $r \in \mathbf{R}^3$.

Il en résulte que *dans un système de coordonnées inertial le second membre de l'équation* (1) *ne peut dépendre que des « coordonnées relatives »* $x_j - x_k$.

L'invariance par le passage à un système de coordonnées se déplaçant uniformément (ce qui ne modifie pas $\ddot{x}_i$ et $x_j - x_k$ mais adjoint à tous les $\dot{x}_j$ un vecteur constant v) entraîne que dans un système de coordonnées inertial le second membre de l'équation (1) est susceptible de dépendre seulement des vitesses relatives

$$\ddot{x}_i = f_i(\{x_j - x_k, \dot{x}_j - \dot{x}_k\}), \quad i, j, k = 1, \ldots, n.$$

Exemple 3. Les transformations galiléennes comprennent les rotations de $\mathbf{R}^3$. L'invariance par ces rotations signifie que cet espace est *isotrope* de sorte qu'il ne possède pas de directions privilégiées.

C'est-à-dire si $\varphi_i: \mathbf{R} \to \mathbf{R}^3$ $(i = 1, \ldots, n)$ est le mouvement d'un système de points vérifiant (1) et $G: \mathbf{R}^3 \to \mathbf{R}^3$ une transformation orthogonale, alors le mouvement $G\varphi_i: \mathbf{R} \to \mathbf{R}^3$ $(i = 1, \ldots, n)$ vérifie également (1). Autrement dit

$$F(Gx, G\dot{x}) = GF(x, \dot{x})$$

où Gx désigne $(Gx_1, \ldots, Gx_n)$, $x_i \in \mathbf{R}^3$.

Exercice. Montrer que *si un système mécanique comprend un point et un seul, alors sa vitesse est nulle dans un système de coordonnées inertial* (« première loi de Newton »).

Indication. En vertu des exemples 1 et 2, le vecteur accélération ne dépend pas de x, $\dot{x}$, t et en vertu de l'exemple 3 le vecteur F est invariant par les rotations.

Exercice. Un système mécanique est composé de deux points. A la date initiale leurs vitesses (dans un système de coordonnées inertial) sont nulles. Montrer que ces points se déplaceront sur la droite qui les joint.

Exercice. Un système mécanique est constitué de trois points. A la date initiale leurs vitesses sont nulles (dans un système de coordonnées inertial). Montrer que ces points ne quitteront pas le plan qui les contenait à la date initiale.

Exercice. Un système mécanique est composé de deux points. Montrer que quelles que soient les conditions initiales il existe un système de coordonnées inertial tel que ces deux points se déplaceront constamment dans un plan fixe.

Exercice. Montrer que dans l'« outre-miroir » la mécanique est identique à la nôtre.

Indication. Le groupe de Galilée renferme des réflexions qui modifient l'orientation de $\mathbf{R}^3$.

Exercice. Etudier l'unicité de la classe des systèmes inertiaux.

Réponse. Cette classe n'est pas unique, les autres s'en déduisent si l'on change l'échelle de longueur et de temps, ainsi que le sens du temps.

§ 3. Exemples de systèmes mécaniques

Nous avons déjà souligné que la forme de la fonction F de l'équation de Newton (1) est définie empiriquement pour chaque système mécanique. Citons quelques exemples.

Lorsqu'on étudie un système concret, il est raisonnable de ne pas y inclure tous les objets de l'Univers. Ainsi, dans la plupart des phénomènes terrestres on ne tient pas compte de l'attraction lunaire. Par ailleurs on peut en général négliger l'influence des processus étudiés sur le mouvement de la Terre et considérer même comme « fixe » le système de coordonnées lié à la Terre. Il va de soi que le principe de relativité n'imposera plus les restrictions précédentes sur les équations de mouvement écrites dans un tel système de coordonnées.

Par exemple, au voisinage de la Terre l'orientation est choisie : c'est celle de la verticale.

A. Exemple 1. Chute libre d'une pierre. L'expérience montre que

$$\ddot{x} = -g, \quad g \approx 9{,}8 \text{ m/s}^2 \text{ (Galilée)}, \tag{2}$$

où x est la hauteur de chute (fig. 6).

Si l'on introduit l'« énergie potentielle » $U = gx$, alors l'équation (2) peut se mettre sous la forme

$$\ddot{x} = -\frac{dU}{dx}.$$

Si $U: E^N \to \mathbf{R}$ est une fonction différentiable dans un espace euclidien, nous désignerons son gradient par $\partial U/\partial \boldsymbol{x}$. Si $E^N = E^{n_1} \times \ldots \times E^{n_k}$ est le produit cartésien d'espaces euclidiens,

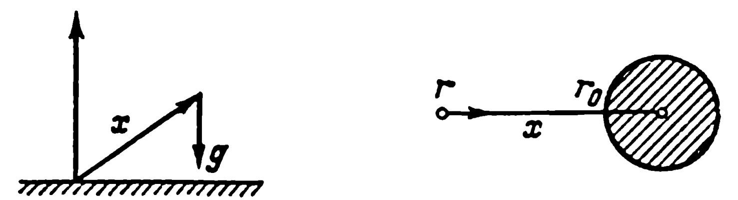


Fig. 6. Chute libre d'une pierre.

Fig. 7. Champ de pesanteur de la Terre.

on désignera le point $\boldsymbol{x} \in E^N$ par $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$ et le vecteur $\partial U/\partial \boldsymbol{x}$ par $(\partial U/\partial \boldsymbol{x}_1, \ldots, \partial U/\partial \boldsymbol{x}_k)$. En particulier, si $x_1, \ldots, x_N$ sont des coordonnées cartésiennes dans E^N, alors les composantes du vecteur $\partial U/\partial \boldsymbol{x}$ seront égales aux dérivées partielles $\partial U/\partial x_1, \ldots, \partial U/\partial x_n$.

L'expérience montre que le rayon vecteur de la pierre relativement à un point quelconque O de la Terre vérifie l'équation

$$\ddot{\boldsymbol{x}} = -\frac{\partial U}{\partial \boldsymbol{x}}, \quad \text{où } U = \boldsymbol{g}\boldsymbol{x}. \tag{3}$$

Le vecteur du second membre est orienté vers la Terre. On l'appelle vecteur accélération de la force de pesanteur $\boldsymbol{g}$.

B. Exemple 2. Chute libre à partir d'un point élevé. Comme tous les faits établis empiriquement, la loi du mouvement (2) est restreinte dans ses applications.

En vertu d'une loi empirique plus précise établie par Newton, l'accélération est *inversement proportionnelle au carré de la distance au centre de la Terre* :

$$\ddot{x} = -g\frac{r_0^2}{r^2},$$

où $r = r_0 + x$ (fig. 7).

On peut donner à cette équation la forme (3) par introduction de l'énergie potentielle

$$U = -\frac{k}{r}, \quad k = gr_0^2,$$

qui est inversement proportionnelle à la distance au centre de la Terre.

Exercice. Déterminer la vitesse qu'il faut communiquer à une pierre pour qu'en quittant la Terre elle se perde à l'infini *).

Réponse. $\geqslant 11{,}2$ km/s.

C. Exemple 3. Mouvement rectiligne d'un poids sous l'action d'un ressort. L'expérience montre que si le ressort s'écarte faiblement

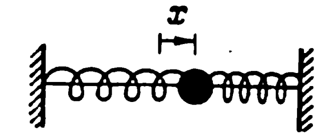


Fig. 8. Poids sur un ressort.

de son état de repos, l'équation du mouvement du poids sera (fig. 8)

$$\ddot{x} = -\alpha^2 x.$$

Cette équation peut se mettre sous la forme (3) par introduction de l'énergie potentielle

$$U = \frac{\alpha^2 x^2}{2}.$$

Si l'on ajoute à ce poids un autre de même masse et que l'on fasse subir au ressort le même allongement, on constate que l'accélération du mouvement est deux fois moindre.

On a établi empiriquement qu'étant donné deux corps quelconques, le rapport $\ddot{x}_1/\ddot{x}_2$ de leurs accélérations est constant pour un même allongement du ressort (ne dépend pas de l'allongement du ressort et de ses propriétés, mais seulement des corps). La grandeur inverse de ce rapport porte le nom de rapport des masses:

$$\frac{\ddot{x}_1}{\ddot{x}_2} = \frac{m_2}{m_1}.$$

Pour unité de masse on adopte la masse d'un corps quelconque fixe, par exemple 1 l d'eau. L'expérience montre que les masses de tous les corps sont positives.

*) C'est ce qu'on appelle la deuxième vitesse cosmique v_2. Notre équation ne tient pas compte de l'attraction du Soleil. La pierre ne quittera pas le système solaire si sa vitesse relativement à la Terre est inférieure à 16,6 km/s.

Le produit $m\ddot{x}$ de la masse d'un corps par son accélération ne dépend pas de ce corps, il caractérise l'allongement du ressort. Cette grandeur est la *force* exercée sur le corps par le ressort.

L'unité de force est le newton. Par exemple, le ressort exerce une force de 9,8 N (= 1 kgf) sur un litre d'eau à la surface du globe.

D. Exemple 4. Système potentiel. Soit $E^{3n} = E^3 \times \ldots \times E^3$ l'espace de configuration d'un système de n points dans un espace euclidien E^3. Soit $U: E^{3n} \to \mathbf{R}$ une fonction différentiable et $m_1, \ldots, m_n$ des nombres positifs.

Définition. *Le mouvement de n points de masses respectives $m_1, \ldots, m_n$ dans un champ potentiel d'énergie potentielle U est régi par le système d'équations différentielles*

$$m_i\ddot{x}_i = -\frac{\partial U}{\partial x_i}, \quad i = 1, \ldots, n. \tag{4}$$

Dans les exemples 1 à 3, les équations de mouvement sont précisément de cette forme. Ont également cette forme les équations de mouvement de la plupart des autres systèmes mécaniques.

Par exemple, on appelle problème de mécanique céleste des trois corps le problème (4) dans lequel

$$U = -\frac{m_1m_2}{\|x_1 - x_2\|} - \frac{m_2m_3}{\|x_2 - x_3\|} - \frac{m_3m_1}{\|x_3 - x_1\|}.$$

A la forme (4) on peut ramener certaines équations différentielles d'une tout autre nature, par exemple les équations des oscillations électriques.

Dans le chapitre suivant nous étudierons essentiellement le système d'équations différentielles (4).

CHAPITRE 2

EQUATIONS DE MOUVEMENT

Dans la plupart des cas (par exemple, dans le problème des trois corps) on n'arrive ni à résoudre le système d'équations différentielles du mouvement ni à étudier assez complètement le comportement des solutions. Dans ce chapitre on se propose de traiter quelques problèmes simples, mais importants faisant intervenir les équations de Newton.

§ 4. Systèmes à un degré de liberté

Dans ce paragraphe on étudiera le plan des phases de l'équation différentielle (1). Pour analyser qualitativement une telle équation il suffit d'un simple coup d'œil sur le graphe de l'énergie potentielle. En outre l'équation (1) est intégrable par quadratures.

A. Définition. Nous appellerons *système à un degré de liberté* un système décrit par une seule équation différentiell

$$\ddot{x} = f(x), \quad x \in \mathbf{R}. \tag{1}$$

L'*énergie cinétique* est la forme quadratique

$$T = \frac{1}{2}\dot{x}^2.$$

L'*énergie potentielle* est la fonction

$$U(x) = -\int_{x_0}^{x} f(\xi)\, d\xi.$$

Dans cette formule le signe est choisi de sorte que plus une pierre sera haut placée, plus son énergie potentielle sera élevée.

A remarquer que l'énergie potentielle U définit f. Donc pour la donnée du système (1) il suffit d'indiquer son énergie potentielle. L'addition d'une constante à l'énergie potentielle ne modifie pas l'équation de mouvement (1).

L'*énergie totale* est la somme

$$E = T + U.$$

Donc l'énergie totale est une fonction $E(x, \dot{x})$.

B. Théorème (loi de conservation de l'énergie). *L'énergie totale d'un point en mouvement* (1) *se conserve*: $E(x(t), \dot{x}(t))$ *ne dépend pas de* t.

Démonstration.

$$\frac{d}{dt}(T+U) = \dot{x}\ddot{x} + \frac{dU}{dx}\dot{x} = \dot{x}(\ddot{x} - f(x)) = 0, \text{ c.q.f.d.}$$

C. Plan des phases.

L'équation (1) équivaut au système de deux équations

$$\dot{x} = y, \quad \dot{y} = f(x). \tag{2}$$

Soit un plan muni des coordonnées x et y. Ce plan est appelé *plan des phases* de l'équation (1). Les points du plan des phases sont appelés *points représentatifs*. Le second membre du système (2) définit un champ de vecteurs sur le plan des phases. Ce champ est appelé *champ de vecteurs de la vitesse de phase.*

Une solution du système (2) est un mouvement $\varphi: \mathbf{R} \to \mathbf{R}^2$ du point représentatif sur le plan des phases, de vitesse égale au vecteur vitesse de phase à l'endroit occupé par le point représentatif à la date considérée *).

L'image de l'application φ est une *courbe de phase.* Donc cette courbe est définie par les équations paramétriques

$$x = \varphi(t), \quad y = \dot{\varphi}(t).$$

Exercice. Montrer que par tout point de l'espace des phases passe une courbe et une seule.

Indication. Voir les cours d'équations différentielles ordinaires.

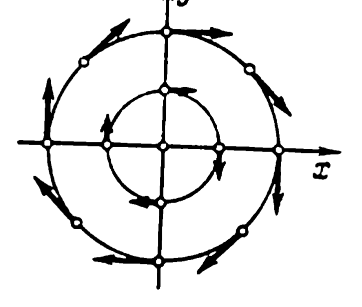


Fig. 9. Plan de phase de l'équation $\ddot{x} = -x$.

A remarquer qu'une courbe de phase peut être constituée seulement d'un point. Un tel point s'appelle *position d'équilibre.* Le vecteur vitesse de phase est nul à la position d'équilibre.

La loi de conservation de l'énergie permet de trouver sans difficultés les courbes de phase. En effet, l'énergie totale a une valeur constante sur chaque courbe de phase. Donc chaque courbe appartient entièrement à un ensemble de niveau d'énergie $E(x, y) = h$.

*) Pour simplifier on suppose que φ est définie sur l'axe temporel $\mathbf{R}$ tout entier.

D. Exemples.

Exemple 1. L'équation fondamentale de la théorie des oscillations s'écrit

$$\ddot{x} = -x.$$

On a dans ce cas (fig. 9):

$$T = \frac{\dot{x}^2}{2}, \quad U = \frac{x^2}{2}, \quad E = \frac{\dot{x}^2}{2} + \frac{x^2}{2}.$$

Les ensembles de niveau d'énergie sont des cercles concentriques et l'origine des coordonnées. En un point (x, y), les composantes du vecteur vitesse de phase sont $(y, -x)$. Il est perpendiculaire et égal en grandeur au rayon vecteur de ce point. Pour cette raison le mouvement du point représentatif sur le plan des phases est une rotation uniforme autour de 0: $x = r_0 \cos(\varphi_0 - t)$, $y = r_0 \sin(\varphi_0 - t)$. Donc tout ensemble de niveau d'énergie est une courbe de phase.

Exemple 2. Supposons que l'énergie potentielle est donnée par son graphe (fig. 10). Traçons les ensembles de niveau d'énergie $\frac{y^2}{2} + U(x) = E$. Ceci étant, il est utile de se rappeler que:

1. Les positions d'équilibre du système (2) sont situées sur l'axe x du plan des phases. Le point $x = \xi$, $y = 0$ est position d'équilibre si ξ est point critique de l'énergie potentielle, i.e. si $dU/dx\,|_{x=\xi} = 0$.

2. Tout ensemble de niveau d'énergie est une courbe *différentiable* au voisinage de chacun de ses points qui ne sont pas position d'équilibre (ceci découle du théorème des fonctions implicites). En particulier, si le nombre E n'est pas valeur critique de l'énergie potentielle (i. e. n'est pas égal à la valeur de l'énergie potentielle en un point critique), alors l'ensemble de niveau où l'énergie est égale à E est une courbe différentiable.

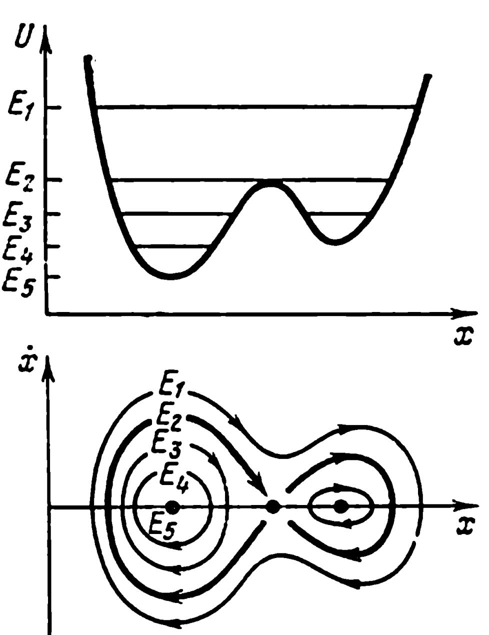


Fig. 10. Energie potentielle et courbes de phase.

Lorsqu'on étudie les lignes de niveau d'énergie, il importe de faire attention aux valeurs critiques E et aux valeurs de E proches des critiques. Pour la commodité il faut imaginer une bille roulant dans une fosse de potentiel U.

Dans le langage préconisé le raisonnement « l'énergie cinétique n'est pas négative. Donc l'énergie potentielle n'est pas supérieure à l'énergie totale. Plus l'énergie potentielle est petite plus la vitesse est élevée » ont pour équivalent « la bille ne peut quitter la fosse

potentielle en atteignant un niveau supérieur à celui défini par son énergie initiale. En roulant dans la fosse, la bille acquiert de la vitesse ». Nous remarquons immédiatement que les points de maximum local de l'énergie potentielle sont des points d'équilibre instable, tandis que les points de minimum, des points d'équilibre stable.

Exercice. Prouver ce qui vient d'être dit.

Exercice. De combien de courbes de phase est constituée la séparatrice, i.e. la courbe (en forme de huit) correspondant au niveau E_2?
Réponse. De trois.

Exercice. Trouver la durée du mouvement sur cette séparatrice.
Réponse. Du théorème d'unicité on déduit que ce temps est infini.

Exercice. Montrer que pour se rendre de x_1 en x_2 (en se déplaçant dans un seul sens) il faut un temps égal à

$$t_2-t_1=\int_{x_1}^{x_2}\frac{dx}{\sqrt{2(E-U(x))}}.$$

Exercice. Déduire le tracé des courbes de phase à partir du graphe de l'énergie potentielle (fig. 11).
Réponse. Fig. 12.

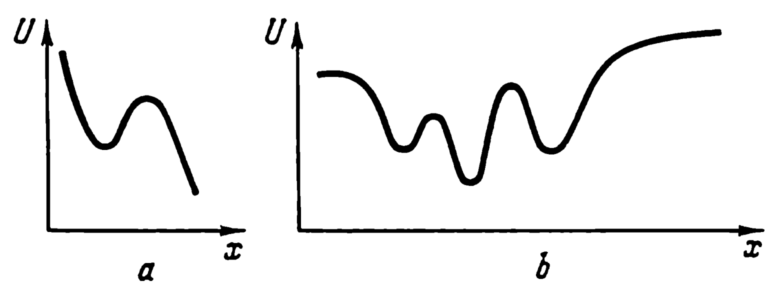


Fig. 11. Energie potentielle.

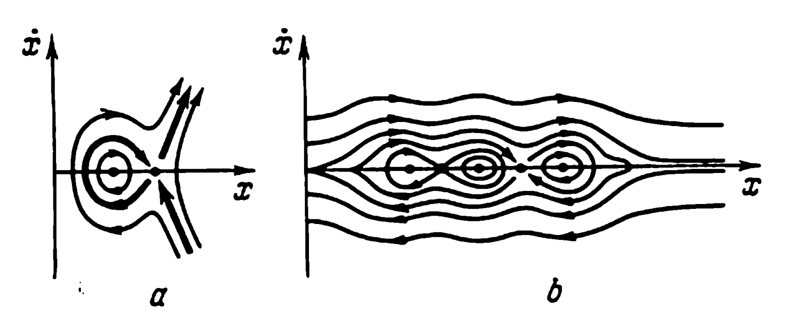


Fig. 12. Courbes de phase.

Exercice. Construire les courbes de phase de l'« équation du pendule mathématique plan »: $\ddot{x} = -\sin x$.

Exercice. Tracer les courbes de phase de l'« équation d'un pendule dont l'axe est en rotation »: $\ddot{x} = -\sin x + M$.

Remarque. Dans ces deux exercices x désigne l'angle d'écart du pendule. Les points représentatifs dont les coordonnées diffèrent de 2π correspondent à une même position du pendule. Aussi, outre le plan des phases est-il naturel de considérer le cylindre des phases $\{x \pmod{2\pi}, y\}$.

Exercice. Trouver les tangentes aux branches de la courbe du niveau critique qui correspond au maximum de l'énergie potentielle $E = U(\xi)$ (fig. 13).

Réponse. $y = \pm\sqrt{U''(\xi)}\,(x - \xi)$.

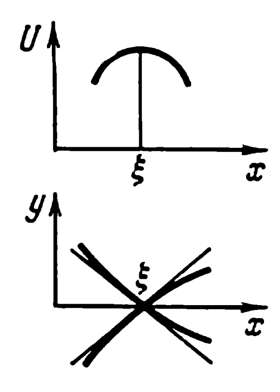


Fig. 13. Ligne de niveau critique d'énergie.

Exercice. Soit $S(E)$ l'aire de la courbe de phase fermée correspondant à un niveau d'énergie E. Montrer que la période du mouvement sur cette courbe est

$$T = \frac{dS}{dE}.$$

Exercice. Soit E_0 la valeur de l'énergie potentielle en un point de minimum ξ_0. Trouver la *période des petites oscillations* au voisinage du point ξ, $T_0 = \lim_{E\to E_0} T(E)$.

Réponse. $2\pi/\sqrt{U''(\xi)}$.

Exercice. Soit un mouvement périodique sur la courbe de phase correspondant au niveau d'énergie E. Est-il stable au sens de Liapounov?

Réponse. Non *).

E. Flot. Soit M un point du plan des phases. Soit une solution du système (2) dont les conditions initiales pour $t = 0$ sont représentées par le point M. Supposons que toute solution du système est prolongeable sur l'axe temporel tout entier. La valeur de cette solution dépend de M pour un certain t. Désignons le point représentatif obtenu (fig. 14) par

$$M(t) = g^t M.$$

Nous avons donc défini une application $g^t: \mathbf{R}^2 \to \mathbf{R}^2$ du plan des phases sur lui-même. D'après les théorèmes connus de la théorie des équations différentielles ordinaires, l'application g^t est un difféomorphisme (i. e. une application bijective et différentiable). Les difféomorphismes g^t, $t \in \mathbf{R}$, forment un groupe : $g^{t+s} = g^t \circ g^s$. L'appli-

*) Sauf si la période ne dépend pas de l'énergie

cation g^0 est l'application identique ($g^0M = M$), et g^{-t}, l'application inverse de g^t. L'application $g: \mathbf{R} \times \mathbf{R}^2 \to \mathbf{R}^2$, $g(t, M) = g^tM$ est différentiable. Toutes ces propriétés traduisent le fait que les

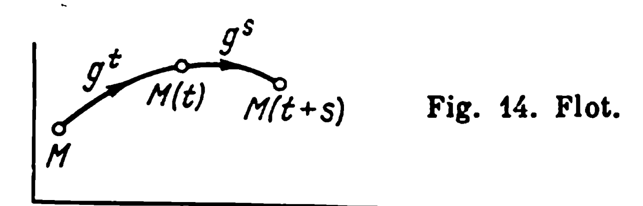


Fig. 14. Flot.

applications g^t forment un *groupe à un paramètre de difféomorphismes* du plan des phases. Ce groupe est également appelé *flot* défini par le système (2) (ou l'équation (1)).

E x e m p l e. Le flot défini par l'équation $\ddot{x} = -x$ est le groupe g^t de rotations du plan des phases d'angle t autour de l'origine des coordonnées.

E x e r c i c e. Montrer qu'un système d'énergie potentielle $U = -x^4$ ne définit aucun flot.

E x e r c i c e. Montrer que si l'énergie potentielle est positive, le flot existe.

I n d i c a t i o n. Se servir de la loi de conservation de l'énergie pour montrer que les solutions sont indéfiniment prolongeables.

E x e r c i c e. Construire l'image du disque $x^2 + (y - 1)^2 < 1/4$ par les

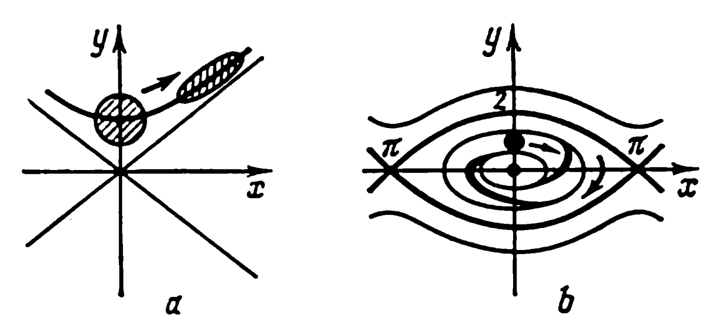


Fig. 15. Effet du flot sur un disque.

applications g^t du flot de l'équation : a) « du pendule renversé » $\ddot{x} = x$, b) « du pendule non linéaire » $\ddot{x} = -\sin x$.

Réponse. Fig. 15.

§ 5. Systèmes à deux degrés de liberté

L'analyse d'un système potentiel général à deux degrés de liberté est hors des possibilités de la science contemporaine. Nous ne verrons que les cas les plus simples dans ce paragraphe.

A. Définition. Un système à deux degrés de liberté est par définition un système décrit par une équation différentielle de la

forme

$$\ddot{x}=f(x), \quad x \in E^2,$$

où f est un champ de vecteurs sur le plan.

Un système est *potentiel* s'il existe une fonction $U: E^2 \to \mathbf{R}$ telle que $f = -\partial U/\partial x$. L'équation du mouvement d'un système potentiel est donc de la forme *)

$$\ddot{x} = -\frac{\partial U}{\partial x}. \tag{1}$$

B. Loi de conservation de l'énergie.

Théorème. *L'énergie totale d'un système potentiel est un invariant:*

$$E=\frac{1}{2}\dot{x}^2+U(x), \quad \dot{x}^2=(\dot{x}, \dot{x}).$$

On affirme que $\frac{dE}{dt}=0$. Démonstration :

$$\frac{dE}{dt}=(\dot{x}, \ddot{x})+\left(\frac{\partial U}{\partial x}, \dot{x}\right)=\left(\ddot{x}+\frac{\partial U}{\partial x}, \dot{x}\right)=0$$

en vertu de l'équation du mouvement.

Corollaire. *Si à la date initiale l'énergie totale est égale à E, la trajectoire tout entière est contenue dans le domaine où $U(x) \leqslant E$*, i. e. le point se trouve tout le temps dans la fosse de potentiel $U(x_1, x_2) \leqslant E$.

Remarque. Dans un système à un degré de liberté on peut toujours introduire l'énergie potentielle

$$U(x) = -\int_{x_0}^{x} f(\xi)\, d\xi.$$

Ce qui n'est pas le cas pour les systèmes à deux degrés de liberté.

Exercice. Citer un exemple de système de la forme $\ddot{x} = f(x)$, $x \in E^2$ qui ne soit pas potentiel.

C. Espace des phases. L'équation de mouvement (1) peut se mettre sous forme du système

$$\left.\begin{array}{ll} \dot{x}_1 = y_1, & \dot{x}_2 = y_2, \\ \dot{y}_1 = -\dfrac{\partial U}{\partial x_1}, & \dot{y}_2 = -\dfrac{\partial U}{\partial x_2}. \end{array}\right\} \tag{2}$$

*) En coordonnées cartésiennes sur E^2:

$$\ddot{x}_1 = -\frac{\partial U}{\partial x_1}, \quad \ddot{x}_2 = -\frac{\partial U}{\partial x_2}.$$

On appelle *espace des phases* d'un système à deux degrés de liberté l'espace de dimension 4 décrit par les coordonnées x_1, x_2, y_1, y_2.

Le système (2) définit le champ de vecteurs de la vitesse de phase sur cet espace des phases et partant *) le flot de notre système (un groupe à un paramètre de difféomorphismes). Les courbes de phase

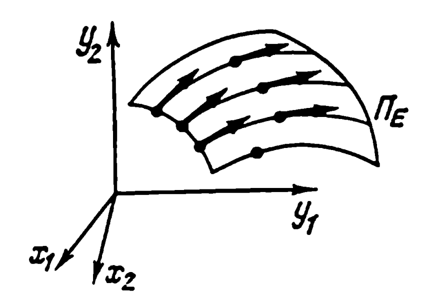


Fig. 16. Surface de niveau d'énergie et courbes de phase.

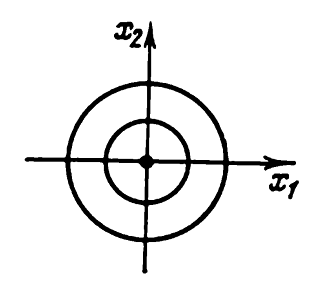


Fig. 17. Lignes de niveau de l'énergie potentielle du pendule sphérique.

du système (2) sont des sous-ensembles de l'espace des phases. L'espace des phases se décompose en courbes de phase. Les projections des courbes de l'espace des phases sur le plan x_1, x_2 donnent les trajectoires du point représentatif sur ce plan. Ces trajectoires sont également appelées orbites. Les orbites peuvent posséder des points doubles contrairement aux courbes de phase. La loi de conservation de l'énergie

$$E = \frac{\dot{x}^2}{2} + U(x) = \frac{y_1^2 + y_2^2}{2} + U(x_1, x_2)$$

définit une hypersurface de dimension 3 dans l'espace de dimension 4 : $E(x_1, x_2, y_1, y_2) = E_0$; cette surface Π_{E_0} est invariante par le flot : $g^t\Pi_{E_0} = \Pi_{E_0}$. Ce qu'on illustre en disant que le flot coule sur la surface de niveau d'énergie. Le champ de vecteurs de la vitesse de phase est tangent à la surface Π_E en chacun de ses points. Donc elle est constituée entièrement de courbes de phase (fig. 16).

D. Exemple 1 (**« petites oscillations du pendule sphérique »**). Soit $U = \frac{x_1^2 + x_2^2}{2}$. Les ensembles de niveau d'énergie potentielle sont des cercles concentriques dans le plan x_1, x_2 (fig. 17).

Les équations du mouvement $\ddot{x}_1 = -x_1$, $\ddot{x}_2 = -x_2$ sont équivalentes au système

$$\dot{x}_1 = y_1, \quad \dot{x}_2 = y_2,$$
$$\dot{y}_1 = -x_1, \quad \dot{y}_2 = -x_2.$$

*) Sous les réserves habituelles, i. e. U est borné inférieurement et deux fois continûment différentiable.

Ce système se décompose en deux systèmes indépendants; autrement dit, chacune des coordonnées x_1, x_2 varie en fonction du temps comme dans un système à un degré de liberté.

Les solutions sont de la forme

$$x_1 = c_1 \cos t + c_2 \sin t, \qquad x_2 = c_3 \cos t + c_4 \sin t,$$
$$y_1 = -c_1 \sin t + c_2 \cos t, \qquad y_2 = -c_3 \sin t + c_4 \cos t.$$

La loi de conservation de l'énergie entraîne

$$E = \frac{1}{2}(y_1^2 + y_2^2) + \frac{1}{2}(x_1^2 + x_2^2) = \text{const},$$

i.e. la surface de niveau Π_E est une sphère dans l'espace des phases $\{(x_1, x_2, y_1, y_2)\}$.

E x e r c i c e. Montrer que les grands cercles de cette sphère sont des courbes de phase. (Un grand cercle est par définition l'intersection de la sphère avec un plan passant par son centre.)

E x e r c i c e. Montrer que l'ensemble des courbes de phase sur la surface Π_E forme une sphère de dimension 2. Plus exactement la formule $w = \frac{x_1 + iy_1}{x_2 + iy_2}$ définit une application de Hopf de la sphère Π_E de dimension 3 sur une sphère de dimension 2 (le plan de la variable complexe w complété du point infini). Nos courbes de phase sont les antécédents de points par l'application de Hopf.

E x e r c i c e. Trouver les projections des courbes de phase sur le plan x_1, x_2 (i.e. tracer les orbites du mouvement des points).

E. Exemple 2 (« figures de Lissajous »). Considérons un autre exemple de mouvement plan (*des « petites oscillations à deux degrés de liberté »*):

$$\ddot{x}_1 = -x_1, \quad \ddot{x}_2 = -\omega^2 x_2.$$

L'énergie potentielle est

$$U = \frac{1}{2} x_1^2 + \frac{1}{2} \omega^2 x_2^2.$$

De la loi de conservation de l'énergie il résulte que si à la date initiale l'énergie totale est

$$\frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + U(x_1, x_2) = E,$$

alors tout le mouvement se déroulera à l'intérieur de l'ellipse $U(x_1, x_2) \leqslant E$.

Par ailleurs, notre système est composé de deux systèmes indépendants de dimension 1. Donc la loi de conservation de l'énergie s'applique à chacun d'eux séparément: les grandeurs

$$E_1 = \frac{1}{2}\dot{x}_1^2 + \frac{1}{2}x_1^2, \qquad E_2 = \frac{1}{2}\dot{x}_2^2 + \frac{1}{2}\omega^2 x_2^2$$
$$(E = E_1 + E_2).$$

se conservent.

Par conséquent, x_1 varie dans la bande $|x_1| \leqslant A_1$, $A_1 = \sqrt{2E_1(0)}$ et x_2 dans la bande $|x_2| \leqslant A_2$. L'intersection de ces deux bandes définit un rectangle qui renferme l'orbite (fig. 18).

E x e r c i c e. Montrer que ce rectangle est inscrit dans l'ellipse $U \leqslant E$.

La solution générale de nos équations est $x_1 = A_1 \sin(t + \varphi_1)$, $x_2 = A_2 \sin(\omega t + \varphi_2)$: le point en mouvement effectue indépendamment des

oscillations de fréquence 1 et d'amplitude A_1 horizontalement, et des oscillations de fréquence ω et d'amplitude A_2 verticalement.

Pour tracer l'orbite sur le plan x_1, x_2 nous allons procéder comme suit. Considérons un cylindre de base $2A_1$ et un ruban de largeur $2A_2$. Décrivons sur

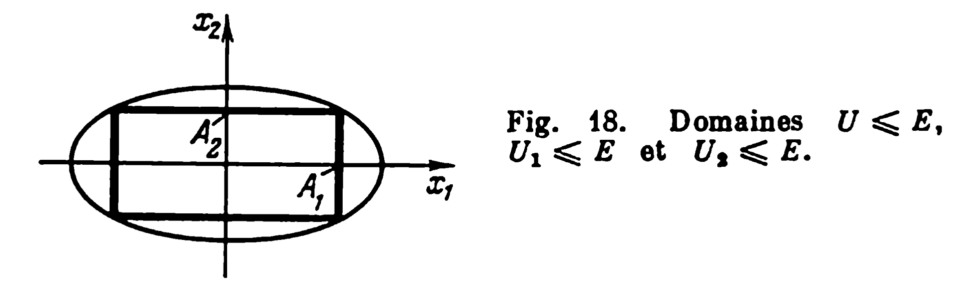


Fig. 18. Domaines $U \leqslant E$, $U_1 \leqslant E$ et $U_2 \leqslant E$.

le ruban une sinusoïde de période $2\pi A_1/\omega$ et d'amplitude A_2 et enroulons-le autour du cylindre (fig. 19). La projection orthogonale de la sinusoïde sur le plan x_1, x_2 nous donnera l'orbite cherchée qu'on appelle *figure de Lissajous*.

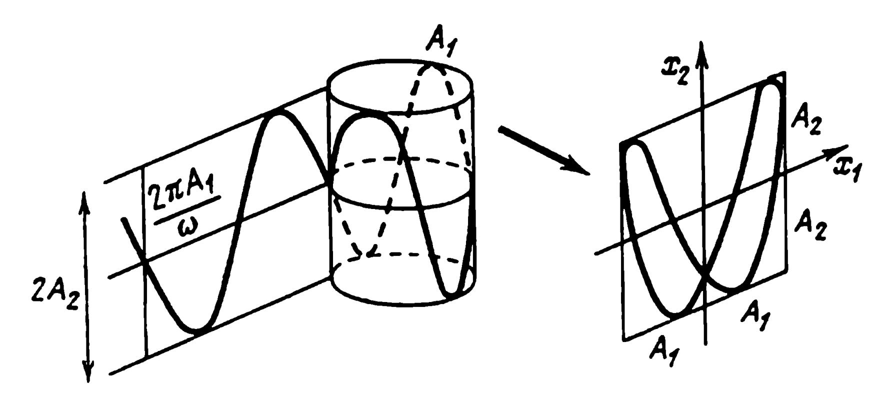


Fig. 19. Construction d'une figure de Lissajous.

Il est commode d'observer les figures de Lissajous sur un oscilloscope en faisant attaquer les plaques de balayage horizontal et vertical par deux oscillations harmoniques indépendantes.

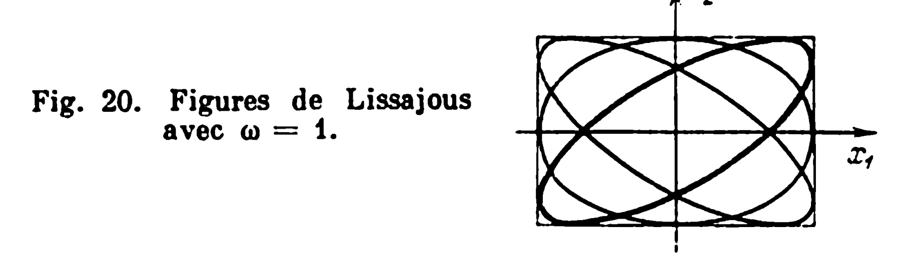


Fig. 20. Figures de Lissajous avec $\omega = 1$.

L'allure affectée par les figures de Lissajous dépend fortement de la fréquence ω. Si $\omega = 1$ (le pendule sphérique de l'exemple 1), on aura une ellipse sur le cylindre. La projection de cette ellipse sur le plan x_1, x_2 dépendra de la différence de phase $\varphi_2 - \varphi_1$. Pour $\varphi_1 = \varphi_2$ on obtient un segment de diagonale du rectangle, pour de faibles $\varphi_2 - \varphi_1$ une ellipse inscrite dans le rectangle et aplatie contre une de ses diagonales. Pour $\varphi_2 - \varphi_1 = \frac{\pi}{2}$ on obtient une ellipse

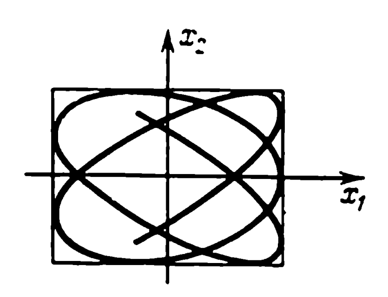


Fig. 21. Figure de Lissajous avec $\omega \approx 1$.

Fig. 22. Figure de Lissajous avec $\omega = 2$.

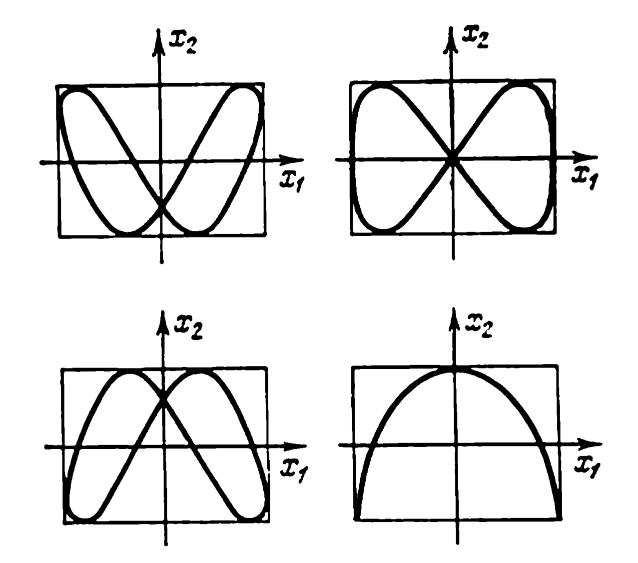


Fig. 23. Figures de Lissajous avec $\omega = 2$.

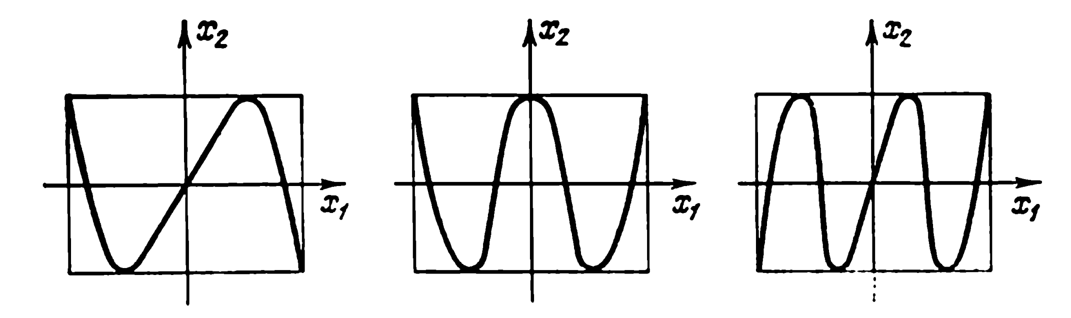


Fig. 24. Polynômes de Tchébychev.

d'axes principaux x_1 et x_2; lorsque $\varphi_2 - \varphi_1$ croît de $\frac{\pi}{2}$ à π, l'ellipse s'aplatit contre l'autre diagonale; lorsque $\varphi_2 - \varphi_1$ continue de croître, ce processus recommence depuis le début (fig. 20).

Supposons maintenant que les fréquences ne sont égales qu'approximativement: $\omega \approx 1$. La portion de courbe correspondant à l'intervalle $0 \leqslant t \leqslant 2\pi$ rappelle beaucoup une ellipse. La spire suivante ressemble aussi à une ellipse, mais elle présente un décalage de phase $\varphi_2 - \varphi_1$ qui est supérieur de $2\pi(\omega - 1)$ à celui de la spire initiale. C'est pourquoi la courbe de Lissajous de fréquence $\omega \approx 1$ est une ellipse qui se déforme en passant par toutes les phases de l'aplatissement contre une diagonale jusqu'à l'aplatissement contre l'autre (fig. 21).

Si une fréquence est le double de l'autre ($\omega = 2$), alors il existe un décalage de phase tel que la figure de Lissajous se transforme en une courbe qui est parcourue deux fois (fig. 22).

E x e r c i c e. Montrer que cette courbe est une parabole.

Lorsque le décalage de phase $\varphi_2 - \varphi_1$ croît, on obtient successivement les courbes de la figure 23.

D'une façon générale, si une fréquence est n fois plus grande que l'autre ($\omega = n$), alors parmi les figures de Lissajous existe le graphe d'un polynôme de degré n (fig. 24) appelé *polynôme de Tchébychev*.

E x e r c i c e. Montrer que si $\omega = m/n$, alors la figure de Lissajous est une courbe algébrique fermée, et si ω est irrationnel, la figure de Lissajous est partout dense dans un rectangle. Trouver l'adhérence de la courbe de phase correspondante.

§ 6. Champ de forces potentiel

Dans ce paragraphe nous allons étudier les liens entre le travail et l'énergie potentielle.

A. Travail d'un champ de forces sur un chemin. Rappelons la définition du travail d'une force $\boldsymbol{F}$ sur un chemin $\boldsymbol{S}$. Le travail d'une

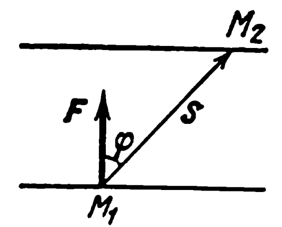


Fig. 25. Travail de la force constante $\boldsymbol{F}$ sur le chemin rectiligne S.

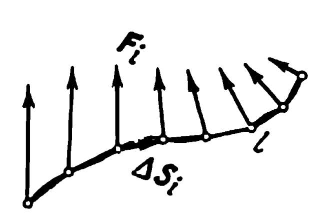


Fig. 26. Travail du champ de forces $\boldsymbol{F}$ sur le chemin l.

force constante $\boldsymbol{F}$ (par exemple, la force avec laquelle nous soulevons un poids) sur un chemin $\boldsymbol{S} = \overrightarrow{M_1M_2}$ est par définition le produit scalaire (fig. 25)

$$A = (\boldsymbol{F}, \boldsymbol{S}) = |\boldsymbol{F}|\,|\boldsymbol{S}|\cdot\cos\varphi.$$

Soient donnés un champ de vecteurs $\boldsymbol{F}$ et une courbe l de longueur finie. Approchons la courbe l par une ligne polygonale de segments

ΔS_i et soit F_i la valeur de la force en un point quelconque de ΔS_i; alors le *travail du champ* F *sur le chemin* l *est par définition* (fig. 26)

$$A = \lim_{|\Delta S_i| \to 0} \sum (F_i, \Delta S_i).$$

En analyse on démontre que si ce champ est continu et le chemin rectifiable, alors cette limite existe et on la note $\int\limits_l (F, dS)$.

B. Conditions de potentialité d'un champ.

Théorème. *Un champ de vecteurs* F *est potentiel si et seulement si son travail sur un chemin quelconque* M_1M_2 *dépend uniquement des extrémités et non de la forme de ce chemin.*

En effet, supposons que le travail du champ F ne dépend pas du chemin. Alors est correctement définie la fonction du point M;

$$U(M) = -\int\limits_{M_0}^{M} (F, dS).$$

On vérifie aisément que

$$F = -\frac{\partial U}{\partial x},$$

i. e. le champ est potentiel et U est son énergie potentielle. De toute évidence, l'énergie potentielle est définie à l'addition près de la constante $U(M_0)$ qui peut être choisie arbitrairement.

Inversement, supposons que le champ F est potentiel et soit U son énergie potentielle. On vérifie alors sans peine que

$$\int\limits_{M_0}^{M} (F, dS) = -U(M) + U(M_0),$$

i. e. le travail ne dépend pas du chemin.

Exercice. Montrer que le champ de vecteurs $F_1 = x_2$, $F_2 = -x_1$ n'est pas potentiel (fig. 27).

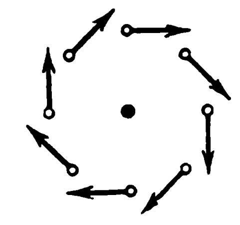


Fig. 27. Champ non potentiel.

Exercice. Etablir si est potentiel le champ $F_1 = \frac{x_2}{x_1^2 + x_2^2}$, $F_2 = \frac{-x_1}{x_1^2 + x_2^2}$ donné sur un plan privé d'un point. Montrer qu'une condition nécessaire et suffisante pour que ce champ soit potentiel est que son travail sur un contour fermé soit nul.

C. Champ central.

D é f i n i t i o n. Un champ de vecteurs sur E^2 est *central* et de centre O s'il est invariant par le groupe des mouvements *) du plan laissant fixe le point O.

E x e r c i c e. Montrer que tous les vecteurs d'un champ central sont portés par des droites passant par O et que le module d'un vecteur dépend seulement de la distance de son point d'application au centre du champ.

Il est utile d'étudier également les champs centraux non définis en O.

E x e m p l e. Le champ newtonien $\boldsymbol{F} = -k\frac{\boldsymbol{r}}{|r|^3}$ est central, ceux des exercices du point B, non.

Théorème. *Tout champ central* $\boldsymbol{F}$ *est potentiel et son énergie potentielle dépend uniquement de la distance à son centre*: $U = U(r)$.

D é m o n s t r a t i o n. En vertu de l'exercice précédent $\boldsymbol{F}(\boldsymbol{r}) = \Phi(r)\,\boldsymbol{e}_r$, où $\boldsymbol{r}$ est le rayon vecteur du point matériel relativement à O, r son module, $\boldsymbol{e}_r$ son vecteur unitaire. On a

$$\int_{M_1}^{M_2} (\boldsymbol{F}\cdot d\boldsymbol{S}) = \int_{r(M_1)}^{r(M_2)} \Phi(r)\,dr,$$

or cette intégrale est visiblement indépendante du chemin.

E x e r c i c e. Calculer l'énergie potentielle du champ newtonien.

R e m a r q u e. Les définitions et théorèmes de ce paragraphe se transposent immédiatement à tout espace euclidien E^n avec n quelconque.

§ 7. Moment cinétique

Dans la suite nous verrons que l'invariance des équations d'un problème mécanique par un groupe quelconque de transformations implique toujours la loi de conservation. Le champ central est invariant par le groupe des rotations. L'intégrale première correspondante porte le nom de moment cinétique.

A. Définition. Le mouvement d'un point matériel (de masse unité) dans un champ central sur un plan est défini par l'équation

$$\ddot{\boldsymbol{r}} = \Phi(r)\,\boldsymbol{e}_r,$$

où $\boldsymbol{r}$ est le rayon vecteur par rapport à O, r son module, $\boldsymbol{e}_r$ son vecteur unitaire. Nous supposerons que notre plan est plongé dans un espace euclidien orienté de dimension 3.

D é f i n i t i o n. *On appelle moment de la quantité de mouvement* (ou *moment cinétique*) d'un point matériel de masse unité,

*) Les réflexions comprises.

relativement au point O, le produit vectoriel

$$\boldsymbol{M} = [\boldsymbol{r}, \dot{\boldsymbol{r}}].$$

Le vecteur $\boldsymbol{M}$ est perpendiculaire au plan et il est défini par un seul nombre : $\boldsymbol{M} = M\boldsymbol{n}$, où $\boldsymbol{n} = [\boldsymbol{e}_1, \boldsymbol{e}_2]$ est le vecteur normal, $\boldsymbol{e}_1$ et $\boldsymbol{e}_2$ le repère orientant le plan (fig. 28).

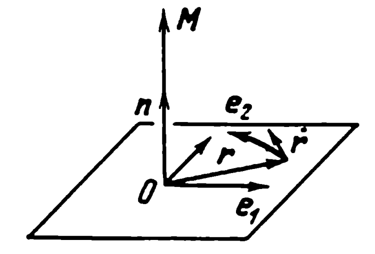


Fig. 28. Moment cinétique.

R e m a r q u e. D'une façon générale, on appelle moment d'un vecteur $\boldsymbol{a}$ appliqué « au point $\boldsymbol{r}$ » relativement à un point O le produit $[\boldsymbol{r}, \boldsymbol{a}]$; dans le cours de statique du secondaire par exemple, on a étudié le moment d'une force.

B. Loi de conservation du moment cinétique.

Lemme. *Soit $\boldsymbol{a}$ et $\boldsymbol{b}$ deux vecteurs variant en fonction du temps dans l'espace euclidien orienté* $\mathbf{R}^3$. *Alors*

$$\frac{d}{dt}[\boldsymbol{a}, \boldsymbol{b}] = [\dot{\boldsymbol{a}}, \boldsymbol{b}] + [\boldsymbol{a}, \dot{\boldsymbol{b}}].$$

D é m o n s t r a t i o n. Ceci découle de la définition de la dérivée.

Théorème (loi de conservation du moment cinétique). *Si le mouvement a lieu dans un champ central, le moment cinétique $\boldsymbol{M}$ relativement au centre O de ce champ est constant.*

D é m o n s t r a t i o n. Par définition

$$\boldsymbol{M} = [\boldsymbol{r}, \dot{\boldsymbol{r}}].$$

En vertu du lemme précédent

$$\dot{\boldsymbol{M}} = [\dot{\boldsymbol{r}}, \dot{\boldsymbol{r}}] + [\boldsymbol{r}, \ddot{\boldsymbol{r}}].$$

De l'équation du mouvement et puisque le champ est central il résulte que les vecteurs $\ddot{\boldsymbol{r}}$ et $\boldsymbol{r}$ sont colinéaires. Donc $\dot{\boldsymbol{M}} = 0$, c.q.f.d.

C. Loi de Kepler. Kepler établit le premier la loi de conservation du moment cinétique en observant le mouvement de Mars. Kepler a énoncé cette loi sous une autre forme.

Introduisons sur le plan les coordonnées polaires r et φ de pôle en O. Considérons au point $\boldsymbol{r}$ de coordonnées $|\boldsymbol{r}| = r$ et φ deux vecteurs unitaires, l'un $\boldsymbol{e}_r$ orienté suivant le rayon vecteur et tel que

$$\boldsymbol{r} = r\boldsymbol{e}_r,$$

l'autre $\boldsymbol{e}_\varphi$, perpendiculaire au premier et orienté dans le sens de

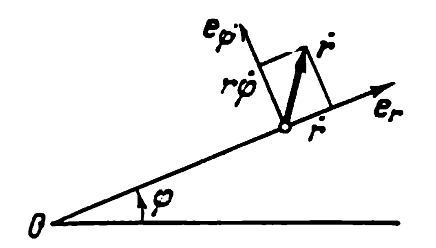


Fig. 29. Développement du vecteur $\dot{\boldsymbol{r}}$ suivant la base $\boldsymbol{e}_r$, $\boldsymbol{e}_\varphi$.

l'accroissement de φ. Décomposons le vecteur vitesse $\dot{\boldsymbol{r}}$ suivant la base $\boldsymbol{e}_r$, $\boldsymbol{e}_\varphi$ (fig. 29).

Lemme. *Est vraie la relation*

$$\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi.$$

Démonstration. De toute évidence, les vecteurs $\boldsymbol{e}_r$, $\boldsymbol{e}_\varphi$ tournent à la vitesse angulaire $\dot{\varphi}$, i.e.

$$\dot{\boldsymbol{e}}_r = \dot{\varphi}\boldsymbol{e}_\varphi,$$

$$\dot{\boldsymbol{e}}_\varphi = -\dot{\varphi}\boldsymbol{e}_r.$$

La dérivation de l'égalité $\boldsymbol{r} = r\boldsymbol{e}_\varphi$ donne

$$\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\boldsymbol{e}}_r = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi, \text{ c.q.f.d.}$$

Donc le moment cinétique est

$$\boldsymbol{M} = [\boldsymbol{r}, \dot{\boldsymbol{r}}] = [\boldsymbol{r}, \dot{r}\boldsymbol{e}_r] + [\boldsymbol{r}, r\dot{\varphi}\boldsymbol{e}_\varphi] = r\dot{\varphi}\,[\boldsymbol{r}, \boldsymbol{e}_\varphi] = r^2\dot{\varphi}\,[\boldsymbol{e}_r, \boldsymbol{e}_\varphi].$$

Ainsi la quantité

$$M = r^2\dot{\varphi}$$

se conserve.

Cette quantité a une intepprétation géométrique simple.

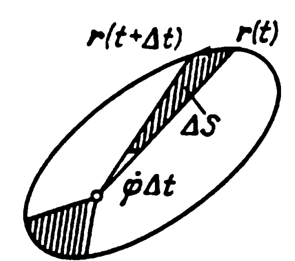
Kepler appela *vitesse aréolaire* C la vitesse de variation de l'aire $S(t)$ balayée par le rayon vecteur (fig. 30):

$$C = \frac{dS}{dt}.$$

Les observations faites par Kepler sur le mouvement des planètes l'ont conduit à énoncer la loi suivante:

Le rayon vecteur balaie des aires égales pendant des intervalles de temps égaux, de sorte que la vitesse aréolaire est constante : $\frac{dS}{dt}$ = const.

Fig. 30. Vitesse aréolaire.



C'est une des formulations de la loi de conservation du moment cinétique. Car

$$\Delta S = S(t+\Delta t) - S(t) = \frac{1}{2} r^2 \dot{\varphi} \Delta t + o(\Delta t),$$

et donc la vitesse aréolaire

$$C = \frac{dS}{dt} = \frac{1}{2} r^2 \dot{\varphi} = \frac{1}{2} M$$

est égale à la moitié du moment cinétique du point matériel de masse unité et, par conséquent, est constante.

Exemple. Les satellites de télécommunications du type « Molnya » ont des orbites fortement allongées. D'après la loi de Kepler, ces satellites passent la plus grande partie de leur temps dans la région éloignée de l'orbite et $\dot{\varphi}$ est petit.

§ 8. Mouvement dans un champ central

La loi de conservation du moment cinétique permet de ramener tout mouvement dans un champ central à un mouvement à un degré de liberté, et partant d'en effectuer une analyse complète.

A. Réduction au problème de dimension 1. Considérons le mouvement d'un point (de masse unité) dans un champ central sur le plan :

$$\ddot{\boldsymbol{r}} = -\frac{\partial U}{\partial \boldsymbol{r}}, \quad U = U(r).$$

Il est naturel de passer aux coordonnées polaires r et φ.

D'après la loi de conservation du moment cinétique la quantité $M = \dot{\varphi}(t)\, r^2(t)$ est constante (ne dépend pas de t).

Théorème. *Si un point matériel de masse unité est en mouvement dans un champ central, sa distance au centre du champ varie comme r dans un problème de dimension 1 d'énergie potentielle*

$$V(r) = U(r) + \frac{M^2}{2r^2}.$$

D é m o n s t r a t i o n. La dérivation de l'expression $\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi$ démontrée au § 5 donne

$$\ddot{\boldsymbol{r}} = (\ddot{r} - r\dot{\varphi}^2)\, \boldsymbol{e}_r + (2\dot{r}\dot{\varphi} + r\ddot{\varphi})\, \boldsymbol{e}_\varphi.$$

Le champ étant central, on a

$$\frac{\partial U}{\partial \boldsymbol{r}} = \frac{\partial U}{\partial r}\, \boldsymbol{e}_r.$$

Donc l'équation du mouvement s'écrit en coordonnées polaires

$$\ddot{r} - r\dot{\varphi}^2 = -\frac{\partial U}{\partial r}, \quad 2\dot{r}\dot{\varphi} + r\ddot{\varphi} = 0.$$

Or en vertu de la loi de conservation du moment cinétique

$$\dot{\varphi} = \frac{M}{r^2},$$

où M est une constante ne dépendant pas de t définie à partir des conditions initiales. Donc

$$\ddot{r} = -\frac{\partial U}{\partial r} + r\frac{M^2}{r^4} \quad \text{ou} \quad \ddot{r} = -\frac{\partial V}{\partial r}, \quad \text{où} \quad V = U + \frac{M^2}{2r^2}.$$

La quantité $V(r)$ est appelée *énergie potentielle effective*.

R e m a r q u e. L'énergie totale

$$E_1 = \frac{\dot{r}^2}{2} + V(r)$$

du problème de dimension 1 obtenu est égale à l'énergie totale

$$E = \frac{\dot{\boldsymbol{r}}^2}{2} + U(r)$$

du problème initial, puisque

$$\frac{\dot{\boldsymbol{r}}^2}{2} = \frac{\dot{r}^2}{2} + \frac{r^2\dot{\varphi}^2}{2} = \frac{\dot{r}^2}{2} + \frac{M^2}{2r^2}.$$

B. Intégration des équations de mouvement. L'énergie totale du problème de dimension 1 obtenu se conserve. Donc on définit r en fonction de t moyennant une quadrature

$$\dot{r} = \sqrt{2(E - V(r))}, \quad \int dt = \int \frac{dr}{\sqrt{2(E - V(r))}}.$$

Comme $\dot{\varphi} = M/r^2$, alors $\dfrac{d\varphi}{dr} = \dfrac{M/r^2}{\sqrt{2(E - V(r))}}$ et l'équation de l'orbite en coordonnées polaires se trouve par une quadrature:

$$\varphi = \int \frac{M/r^2\, dr}{\sqrt{2(E - V(r))}}.$$

C. Orbites. Fixons le moment constant M. On étudie facilement les variations de r en fonction du temps en traçant le graphe de l'énergie potentielle effective $V(r)$ (fig. 31).

Soit E la valeur de l'énergie totale. L'orbite correspondant à ces E et M est entièrement contenue dans le domaine $V(r) \leqslant E$. Sur

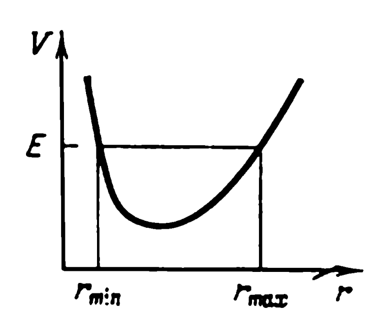


Fig. 31. Graphe de l'énergie potentielle effective.

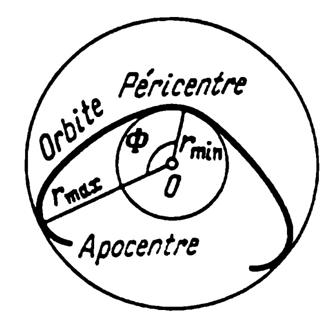


Fig. 32. Orbite du point dans un champ central.

la frontière de ce domaine $V = E$, i.e. $\dot{r} = 0$. La vitesse du point en mouvement n'est en général pas nulle, puisque $\dot{\varphi} \neq 0$ pour $M \neq 0$.

L'inégalité $V(r) \leqslant E$ définit sur le plan une ou plusieurs couronnes:

$$0 \leqslant r_{min} \leqslant r \leqslant r_{max} \leqslant \infty.$$

Si $0 < r_{min} < r_{max} < \infty$, le mouvement est borné et se déroule à l'intérieur de la couronne déterminée par les cercles de rayons r_{min} et r_{max}.

L'orbite est représentée sur la figure 32. L'angle φ varie monotonement, r oscille périodiquement entre r_{min} et r_{max}. Les points tels que $r = r_{min}$ sont appelés *péricentres* et tels que $r = r_{max}$, *apocentres* (resp. périgée et apogée si le centre est la Terre, périhélie et aphélie si c'est le Soleil et périsélénie et aposélénie si c'est la Lune).

Tout axe joignant le centre à un apocentre ou à un péricentre est axe de symétrie de l'orbite.

En général, une orbite n'est pas fermée: l'angle d'un péricentre et d'un apocentre successifs est donné par l'intégrale

$$\Phi = \int_{r_{min}}^{r_{max}} \frac{M/r^2\, dr}{\sqrt{2(E - V(r))}}.$$

L'angle de deux péricentres successifs est égal à 2Φ.

Une orbite est fermée si l'angle Φ est commensurable avec 2π, i. e. si $\Phi = 2\pi \frac{m}{n}$, où m et n sont des entiers.

On montre que si Φ est incommensurable avec 2π, alors l'orbite est partout dense dans une couronne (fig. 33).

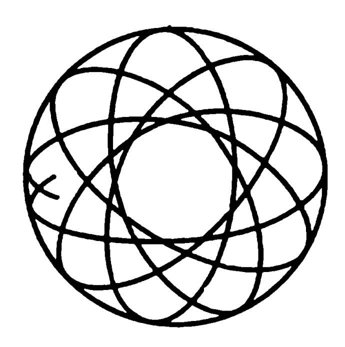


Fig. 33. Orbite partout dense dans une couronne.

Si $r_{\min} = r_{\max}$, i.e.[1] E est la valeur prise par V en un point de minimum, alors la couronne dégénère en un cercle qui sera précisément orbite.

E x e r c i c e. Pour quelles valeurs de α le mouvement suivant une orbite circulaire dans un champ d'énergie potentielle $U = r^{\alpha}$, $-2 \leqslant \alpha < \infty$, est-il stable au sens de Liapounov?
Réponse. Uniquement pour $\alpha = 2$.

Pour des valeurs de E légèrement supérieures au minimum de V la couronne $r_{\min} \leqslant r \leqslant r_{\max}$ sera très étroite; quant à l'orbite, elle sera proche d'un cercle. Dans le problème de dimension 1 considéré, r effectuera des petites oscillations au voisinage d'un point de minimum de V.

E x e r c i c e. Calculer Φ pour une orbite proche d'une orbite circulaire de rayon r.
I n d i c a t i o n. Voir le point D plus bas.
Traitons maintenant le cas $r_{\max} = \infty$. Si $\lim\limits_{r\to\infty} U(r) = \lim\limits_{r\to\infty} V(r) = U_{\infty} < \infty$, il est possible que le point s'éloigne à l'infini. Si l'énergie initiale E est supérieure à U_{∞}, alors le point s'éloigne à l'infini avec une vitesse finie $\dot{r}_{\infty} = \sqrt{2(E - U_{\infty})}$. On remarquera que si $U(r)$ tend vers sa limite moins vite que r^{-2}, alors le potentiel effectif V sera attractif à l'infini (on suppose ici que le potentiel U est attractif à l'infini).
Si pour $r \to 0$ $|U(r)|$ ne croît pas plus vite que $M^2/2r^2$, alors $r_{\min} > 0$ et l'orbite ne s'approche pas du centre. Si $U(r) + M^2/2r^2 \to -\infty$ pour $r \to 0$, alors est possible une « chute au centre du champ ». On peut atteindre le centre du champ même dans un espace de temps fini (par exemple dans le champ $U(r) = -1/r^3$).

E x e r c i c e. Etudier l'allure des orbites lorsque l'énergie totale est égale à la valeur de l'énergie effective en un point de maximum local.

D. Champs centraux dans lesquels toutes les orbites bornées sont fermées. Des exercices qui vont suivre il découle que dans un champ central les orbites bornées sont fermées dans deux cas seulement:

$$U = ar^2, \quad a \geqslant 0 \quad \text{et} \quad U = -k/r, \quad k \geqslant 0.$$

Exercice 1. Montrer que l'angle Φ d'un péricentre et d'un apocentre successifs est égal à une demi-période des oscillations dans un système de dimension 1 d'énergie potentielle $W(x) = U\left(\frac{M}{x}\right) + \frac{x^2}{2}$.

Indication. La substitution $x = M/r$ donne

$$\Phi = \int_{x_{\min}}^{x_{\max}} \frac{dx}{\sqrt{2(E-W)}}.$$

Exercice 2. Trouver l'angle Φ pour une orbite proche d'une orbite circulaire de rayon r.

Réponse. $\Phi \approx \Phi_{\text{cir}} = \pi \dfrac{M}{r^2\sqrt{W''(r)}} = \pi \sqrt{\dfrac{U'}{3U' + rU''}}$.

Exercice 3. Pour quels U la quantité Φ_{cir} ne dépend-elle pas du rayon r?

Réponse. $U(r) = ar^\alpha$ $(\alpha \geqslant -2, \alpha \neq 0)$ et $U(r) = b \log r$.

Ici $\Phi_{\text{cir}} = \pi/\sqrt{\alpha + 2}$ (le cas logarithmique correspond à $\alpha = 0$). Par exemple, pour $\alpha = 2$, on a $\Phi_{\text{cir}} = \pi/2$ et pour $\alpha = -1$ $\Phi_{\text{cir}} = \pi$.

Exercice 4. Soit $U(r) \to \infty$ pour $r \to \infty$. Trouver $\lim\limits_{E\to\infty} \Phi(E, M)$.

Réponse. $\pi/2$.

Indication. La substitution $x = yx_{\max}$ ramène Φ à la forme

$$\Phi = \int_{y_{\min}}^{1} \frac{dy}{\sqrt{2(W^*(1) - W^*(y))}}, \quad W^*(y) = \frac{y^2}{2} + \frac{1}{x_{\max}^2} U\left(\frac{M}{yx_{\max}}\right).$$

Lorsque $E \to \infty$, on a $x_{\max} \to \infty$, $y_{\min} \to 0$ et l'on peut éliminer le second terme de W^*.

Exercice 5. Soit $U(r) = -kr^{-\beta}$, $0 < \beta < 2$. Trouver $\Phi_0 = \lim\limits_{E\to-0} \Phi$.

Réponse. $\Phi_0 = \displaystyle\int_0^1 \frac{dx}{\sqrt{x^\beta - x^2}} = \frac{\pi}{2-\beta}$. On remarquera que Φ_0 ne dépend pas de M.

Exercice 6. Trouver tous les champs centraux dont les orbites bornées existent et sont toutes fermées.

Réponse. $U = ar^2$ ou $U = -k/r$.

Solution. Si toutes les orbites bornées sont fermées, on a en particulier $\Phi_{\text{cir}} = 2\pi \dfrac{m}{n} = \text{const}$. En vertu de l'exercice 3 soit $U = ar^\alpha$ $(\alpha \geqslant -2)$, soit $U = b \ln r$ $(\alpha = 0)$. Dans les deux cas $\Phi_{\text{cir}} = \pi/\sqrt{\alpha + 2}$. Si $\alpha > 0$,

de l'exercice 4 il vient $\lim_{E\to\infty} \Phi(E, M) = \pi/2$. Donc $\Phi_{\text{cir}} = \pi/2$, $\alpha = 2$. Si $\alpha < 0$, alors en vertu de l'exercice 5, $\lim_{E\to -0} \Phi(E, M) = \pi/(2+\alpha)$. Par conséquent $\pi/(2+\alpha) = \pi/\sqrt{2+\alpha}$, $\alpha = -1$. Pour $\alpha = 0$ on trouve $\Phi_{\text{cir}} = \pi/\sqrt{2}$, valeur qui est incommensurable avec 2π. Donc toutes les obrites bornées ne peuvent être fermées que dans les champs $U = ar^2$ ou $U = -k/r$. Dans le champ $U = ar^2$, $a > 0$, toutes les orbites sont fermées (ce sont des ellipses de centre O; cf. exemple 1 du § 3). Dans le champ $U = -k/r$ toutes les orbites bornées sont également fermées et également elliptiques comme nous allons le montrer.

E. Le problème de Kepler. Il s'agit du mouvement dans un champ central de potentiel $U = -k/r$ et donc $V(r) = -k/r + M^2/2r^2$ (fig. 34).

D'après la formule générale

$$\varphi = \int \frac{M/r^2\, dr}{\sqrt{2(E - V(r))}}.$$

On obtient par intégration

$$\varphi = \arccos \frac{\frac{M}{r} - \frac{k}{M}}{\sqrt{2E + \frac{k^2}{M^2}}}.$$

En principe il faudrait ajouter à cette expression une constante arbitraire. Mais nous la supposons nulle, ce qui revient à compter

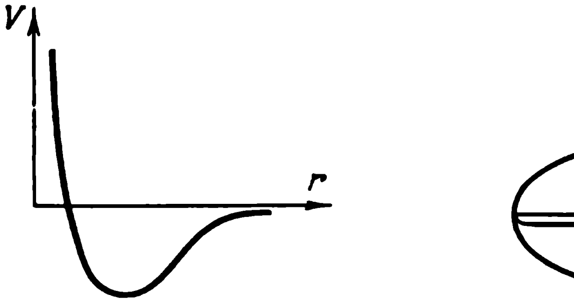


Fig. 34. Potentiel effectif du problème de Kepler.

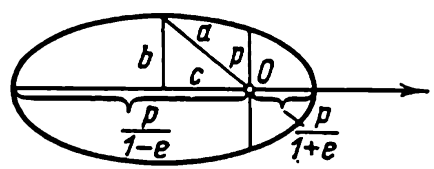


Fig. 35. Ellipse keplérienne.

l'angle φ à partir du péricentre. Introduisons les notations suivantes:

$$\frac{M^2}{k} = p, \quad \sqrt{1 + \frac{2EM^2}{k^2}} = e.$$

Il vient alors $\varphi = \arccos \frac{p/r - 1}{e}$, i.e.

$$r = \frac{p}{1 + e\cos\varphi}.$$

C'est ce qu'on appelle l'*équation focale* de la section conique. Le mouvement est borné (fig. 35) pour $E < 0$. Alors $e < 1$, i. e. la section conique est une ellipse. La quantité p est le *paramètre* de l'ellipse, et e, l'*excentricité*. La première loi que Kepler a découverte en observant le mouvement de Mars dit que les planètes décrivent des ellipses dont le Soleil occupe l'un des foyers.

Si l'on admet que les planètes sont en mouvement dans un champ central de gravitation, alors la première loi de Kepler entraîne la loi de l'attraction universelle de Newton: $U = -k/r$ (cf. point D ci-dessus).

Le paramètre et l'excentricité sont liés aux demi-axes par les relations

$$2a = \frac{p}{1-e} + \frac{p}{1+e} = \frac{2p}{1-e^2}, \quad \text{i.e. } a = \frac{p}{1-e^2};$$

$e = \frac{c}{a} = \frac{\sqrt{a^2-b^2}}{a}$, où $c = ae$ est la distance du centre à un foyer (voir fig. 35).

R e m a r q u e. Une ellipse de faible excentricité ressemble beaucoup à un cercle *). Si la distance d'un foyer au centre est du premier ordre de petitesse, alors la différence des demi-axes est du second ordre: $b = a\sqrt{1-e^2} \approx a\left(1 - \frac{e^2}{2}\right)$. Exemple: dans une ellipse de grand axe 10 cm et d'excentricité 0,1, la différence des demi-axes est 0,5 mm et la distance d'un foyer au centre 1 cm.

L'excentricité des orbites des planètes est très petite. Ceci explique que Kepler ait d'abord énoncé sa première loi comme suit: les planètes se déplacent autour du Soleil suivant des cercles dont le Soleil n'occupe pas le centre.

La deuxième loi de Kepler: la vitesse aréolaire est constante, est valable dans tout champ central.

La troisième loi de Kepler s'énonce comme suit: la durée d'une révolution sur une orbite elliptique dépend uniquement de la longueur du grand axe.

Les carrés des périodes de révolution de deux planètes en mouvement sur deux orbites elliptiques différentes sont entre eux comme les cubes des grands axes **).

D é m o n s t r a t i o n. Désignons par T la période de révolution, et par S l'aire balayée par le rayon vecteur dans le temps T. On a $2S = MT$, puisque $M/2$ est la vitesse aréolaire. Or l'aire de l'ellipse est égale à $S = \pi ab$, donc $T = \frac{2\pi ab}{M}$. Comme

*) Laissez tomber une goutte de thé non loin du centre de la tasse. Les ondes se rassembleront au point symétrique. La raison est qu'en vertu de la définition focale de l'ellipse, les ondes issues d'un foyer de l'ellipse se rassemblent dans l'autre.

**) Les planètes sont considérées ici comme des points en mouvement dans un champ central.

$$=\frac{M^2/k}{2\,|E|\,\frac{M^2}{k^2}}=\frac{k}{2\,|E|}\left(a=\frac{p}{1-e^2}\right),\quad b=\frac{M^2}{k}\cdot\frac{1}{\sqrt{2\,|E|}\cdot\frac{M}{k}}=\frac{M}{\sqrt{2\,|E|}},$$

alors $T=2\pi\frac{k}{(\sqrt{2\,|E|})^3}$; or $2\,|E|=\frac{k}{a}$, donc $T=2\pi a^{3/2}k^{-1/2}$.

On remarquera que l'énergie totale E dépend uniquement du demi grand axe a de l'orbite et qu'elle est la même pour toute la famille d'orbites elliptiques (du cercle de rayon a au segment de longueur $2a$).

E x e r c i c e. Pendant la mise sur orbite circulaire d'un satellite à 300 km de la Terre, la direction de la vitesse s'est écartée de la direction théorique de 1° dans le sens de la Terre. Comment variera le périgée?

Réponse. La hauteur du périgée diminuera d'environ 110 km.

I n d i c a t i o n. La différence entre cette orbite et un cercle est du deuxième ordre de petitesse et on peut la négliger. Le rayon est égal au rayon théorique, puisque l'énergie initiale est égale à l'énergie initiale théorique. Donc l'orbite se déduit de l'orbite théorique par une rotation de 1° (fig. 36).

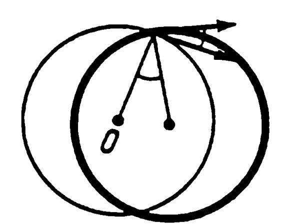


Fig. 36. Orbite proche d'une orbite circulaire.

E x e r c i c e. Comment variera la hauteur du périgée si la vitesse sera de 1 m/s inférieure à la vitesse théorique?

E x e r c i c e. On appelle *première vitesse cosmique* la vitesse du mouvement suivant une orbite circulaire de rayon proche de celui de la Terre. Trouver la première vitesse cosmique v_1 et montrer que $v_2=\sqrt{2}v_1$ (cf. § 3, B).

Réponse. 8,1 km/s.

E x e r c i c e *). Pendant sa « promenade » dans l'espace, le cosmonaute A. Léonov a lancé en direction de la Terre l'obturateur de sa caméra. Etudier le mouvement de cet obturateur par rapport au vaisseau cosmique en supposant qu'il a été lancé à 10 m/s.

Réponse. L'obturateur se déplacera par rapport au cosmonaute suivant une ellipse dont le grand axe vaut environ 32 km et le petit axe environ 16 km. Le centre de l'ellipse est situé à 16 km en avant sur l'orbite, la période de révolution sur l'ellipse est égale à la période du mouvement sur l'orbite.

I n d i c a t i o n. Prenons pour unité de longueur le rayon de l'orbite circulaire du vaisseau cosmique et choisissons l'unité de temps telle que la période de révolution sur cette orbite soit égale à 2π. Il nous faut étudier les solutions de l'équation de Newton

$$\ddot{\mathbf{r}}=-\mathbf{r}/r^3,$$

proches de la solution circulaire $r_0=1$, $\varphi_0=t$. Cherchons cette solution sous la forme

$$r=r_0+r_1,\quad \varphi=\varphi_0+\varphi_1\qquad r_1\ll 1,\quad \varphi_1\ll 1.$$

*) Ce problème a été emprunté à l'attrayant livre de V. B é l e t s k i, *Sur le mouvement des engins spatiaux* (en russe), « Naouka », Moscou, 1972.

D'après le théorème de différentiabilité de la solution par rapport aux conditions initiales, les fonctions $r_1(t)$ et $\varphi_1(t)$ vérifient un système d'équations différentielles linéaires (équations aux variations) aux infiniment petits du premier ordre près de l'écart initial.

En portant les expressions de r et φ dans l'équation de Newton, on obtient, tous calculs faits, les équations aux variations

$$\ddot{r}_1 = 3r_1 + 2\dot{\varphi}_1, \quad \ddot{\varphi}_1 = -2\dot{r}_1.$$

La résolution de ces équations avec les conditions initiales ($r_1(0) = \varphi_1(0) = \dot{\varphi}_1(0) = 0$, $\dot{r}_1(0) = -1/800$) nous conduit à la réponse donnée plus haut.

Les infiniment petits du deuxième ordre que nous avons négligés représentent environ le 1/800 du résultat obtenu (i. e. environ quelques dizaines de mètres par révolution). Donc après avoir parcouru une ellipse de 30 km en une heure et demie, l'obturateur revient vers le vaisseau cosmique du côté opposé à la Terre et passe à quelques dizaines de mètres de lui.

Il est évident que dans nos calculs nous avons négligé le fait que l'orbite n'était pas circulaire, l'influence des forces autres que la force de gravitation, etc.

§ 9. Mouvement d'un point dans un espace de dimension 3

Dans ce paragraphe nous définirons le moment cinétique par rapport à un axe et démontrerons que ce moment se conserve si le mouvement se déroule dans un champ à symétrie axiale.

Tous les résultats établis pour le mouvement d'un point dans le plan s'étendent sans difficultés au mouvement dans l'espace.

A. Champ potentiel. Soit le mouvement d'un point matériel dans un champ potentiel

$$\ddot{\boldsymbol{r}} = -\frac{\partial U}{\partial \boldsymbol{r}}, \quad \text{où } U = U(\boldsymbol{r}), \quad \boldsymbol{r} \in E^3.$$

La *loi de conservation de l'énergie* est vérifiée

$$\frac{dE}{dt} = 0 \quad \left(\text{où } E = \frac{1}{2}\dot{\boldsymbol{r}}^2 + U(\boldsymbol{r})\right).$$

B. Champ central. *La loi de conservation du moment de la quantité de mouvement est* $\boldsymbol{M} = [\boldsymbol{r}, \dot{\boldsymbol{r}}]$. *Le vecteur* $\boldsymbol{M}$ *est invariant si le mouvement s'effectue dans un champ central*, i.e.

$$\frac{d\boldsymbol{M}}{dt} = 0.$$

Tout champ central est potentiel (la démonstration se fait comme pour le cas bidimensionnel) et

$$\frac{d\boldsymbol{M}}{dt} = [\dot{\boldsymbol{r}}, \dot{\boldsymbol{r}}] + [\boldsymbol{r}, \ddot{\boldsymbol{r}}] = 0,$$

puisque $\ddot{\boldsymbol{r}} = -\partial U/\partial \boldsymbol{r}$ et les vecteurs $\partial U/\partial \boldsymbol{r}$ et $\boldsymbol{r}$ sont colinéaires, car le champ est central.

Corollaire. *Toute orbite est plane si le mouvement a lieu dans un champ central.*

Démonstration. $(\boldsymbol{M}, \boldsymbol{r}) = ([\boldsymbol{r}, \dot{\boldsymbol{r}}], \boldsymbol{r}) = 0$, donc $\boldsymbol{r}(t) \perp \boldsymbol{M}$ et comme $\boldsymbol{M} = \text{const}$, l'orbite tout entière est contenue dans un plan orthogonal à $\boldsymbol{M}$ *). Donc l'étude d'une orbite dans un champ central dans l'espace se ramène au problème plan traité dans le paragraphe précédent.

Exercice. Etudier le mouvement d'un point dans un champ central dans un espace euclidien de dimension n.

C. Champ à symétrie axiale. Définition. Un champ de vecteurs sur E^3 est *à symétrie axiale* s'il est invariant par le groupe des rotations de l'espace laissant fixe chaque point d'un certain axe.

Exercice. Montrer que si un champ est à symétrie axiale et potentiel, son énergie potentielle est de la forme $U = U(r, z)$, où r, φ, z sont les coordonnées cylindriques.

En particulier, il en découle que le vecteur du champ est situé dans un plan contenant la droite z.

Un tel champ peut être le champ de gravitation engendré par un corps de révolution.

Soient z une droite orientée par le vecteur unitaire $\boldsymbol{e}_z$ dans l'espace euclidien orienté E^3, $\boldsymbol{F}$ un vecteur de l'espace vectoriel euclidien

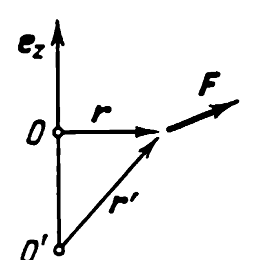


Fig. 37. Moment du vecteur $\boldsymbol{F}$ par rapport à l'axe z.

$\mathbf{R}^3$, O un point de z, $\boldsymbol{r} = x - O \in \mathbf{R}^3$ le rayon vecteur du point $x \in E^3$ par rapport à O (fig. 37).

Définition. On appelle *moment M_z du vecteur $\boldsymbol{F}$, appliqué au point $\boldsymbol{r}$ par rapport à z*, la projection sur cette droite du moment du vecteur $\boldsymbol{F}$ par rapport à un point quelconque de la droite z:

$$M_z = (\boldsymbol{e}_z, [\boldsymbol{r}, \boldsymbol{F}]).$$

Le nombre M_z ne dépend pas du choix du point O sur la droite z. En effet, soit un point O' de cette droite. D'après la propriété du produit mixte: $M_z' = (\boldsymbol{e}_z, [\boldsymbol{r}', \boldsymbol{F}]) = ([\boldsymbol{e}_z, \boldsymbol{r}'], \boldsymbol{F}) = ([\boldsymbol{e}_z, \boldsymbol{r}], \boldsymbol{F}) = M_z$.

Remarque. M_z dépend du choix de l'*orientation* de z: si l'on remplace $\boldsymbol{e}_z$ par $-\boldsymbol{e}_z$, alors M_z change de signe.

*) Le cas $\boldsymbol{M} = 0$ est laissé au soin du lecteur.

Théorème. *Si le mouvement s'effectue dans un champ potentiel à symétrie axiale autour de la droite z, le moment de la quantité de mouvement relativement à z se conserve.*

Démonstration. $M_z = (\boldsymbol{e}_z, [\boldsymbol{r}, \dot{\boldsymbol{r}}])$,

$$\dot{M}_z = (\boldsymbol{e}_z, [\dot{\boldsymbol{r}}, \dot{\boldsymbol{r}}]) + (\boldsymbol{e}_z, [\boldsymbol{r}, \ddot{\boldsymbol{r}}]) = 0,$$

comme $\ddot{\boldsymbol{r}} = \boldsymbol{F}$, les vecteurs $\boldsymbol{r}$ et $\ddot{\boldsymbol{r}}$ sont situés dans un plan contenant la droite z et donc $[\boldsymbol{r}, \ddot{\boldsymbol{r}}]$ est perpendiculaire à $\boldsymbol{e}_z$.

Remarque. La démonstration reste en vigueur pour tout champ de forces dont le vecteur force $\boldsymbol{F}$ est contenu dans le plan engendré par $\boldsymbol{r}$ et $\boldsymbol{e}_z$.

§ 10. Mouvement d'un système de n points

Dans ce paragraphe on démontre les lois de conservation de l'énergie, de l'impulsion et du moment cinétique pour un système de points matériels dans E^3.

A. Forces intérieures et extérieures. On appelle équations de Newton du mouvement d'un système de n points matériels de masses m_i et de rayons vecteurs $\boldsymbol{r}_i \in E^3$ les équations

$$m_i \boldsymbol{r}_i = \boldsymbol{F}_i, \quad i = 1, 2, \ldots, n.$$

Le vecteur $\boldsymbol{F}_i$ est appelé *force agissant sur le i-ème point.*

Les forces $\boldsymbol{F}_i$ sont définies empiriquement. Les observations montrent que souvent dans un système de deux points ces forces

Fig. 38. Forces d'interaction.

sont égales en grandeur, de sens contraires et leurs actions sont dirigées le long de la droite qui joint ces points (fig. 38).

De telles forces sont appelées *forces d'interaction* (exemple : les forces de gravitation universelle).

Si toutes les forces agissant sur les points d'un système sont des forces d'interaction, alors ce système est dit *fermé*. Dans un système fermé la force qui agit sur le i-ème point est par définition

$$\boldsymbol{F}_i = \sum_{\substack{j=1 \\ i \neq j}}^{n} \boldsymbol{F}_{ij}.$$

Le vecteur $\boldsymbol{F}_{ij}$ est la force avec laquelle le j-ème point agit sur le i-ème.

Comme les forces $\boldsymbol{F}_{ij}$ et $\boldsymbol{F}_{ji}$ sont opposées ($\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$), on peut les mettre sous la forme $\boldsymbol{F}_{ij} = f_{ij}\boldsymbol{e}_{ij}$, où $f_{ij} = f_{ji}$ est la grandeur de la force et $\boldsymbol{e}_{ij}$ le vecteur unitaire de la droite passant par les points i et j dans le sens de i à j.

Si le système n'est pas fermé, souvent on peut représenter les forces qui le sollicitent sous la forme

$$\boldsymbol{F}_i = \sum \boldsymbol{F}_{ij} + \boldsymbol{F}'_i,$$

où $\boldsymbol{F}_{ij}$ sont les forces d'interaction et $\boldsymbol{F}'_i(\boldsymbol{r}_i)$ une force qu'on appelle *force extérieure.*

E x e m p l e (fig. 39). Partageons le système fermé en deux parties I et II. La force $\boldsymbol{F}_i$ appliquée au i-ème point du système I

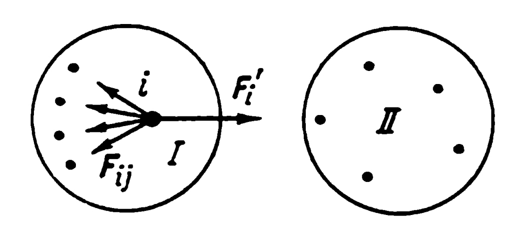


Fig. 39. Forces intérieures et extérieures.

est définie par les forces d'interaction à l'intérieur du système I et les forces exercées sur le i-ème point par les points du système II, i. e.

$$\boldsymbol{F}_i = \sum_{\substack{j \in I \\ i \neq j}} \boldsymbol{F}_{ij} + \boldsymbol{F}'_i.$$

$\boldsymbol{F}'_i$ représente une force extérieure au système I.

B. Loi de conservation de l'impulsion.

D é f i n i t i o n. On appelle *impulsion* (ou *quantité de mouvement*) d'un système le vecteur

$$\boldsymbol{P} = \sum_{i=1}^{n} m_i \dot{\boldsymbol{r}}_i.$$

Théorème. *La vitesse de variation de la quantité de mouvement d'un système est égale à la somme de toutes les forces extérieures agissant sur les points de ce système.*

D é m o n s t r a t i o n.

$$\frac{d\boldsymbol{P}}{dt} = \sum_{i=1}^{n} m_i \ddot{\boldsymbol{r}}_i = \sum_{i=1}^{n} \boldsymbol{F}_i = \sum_{i,j} \boldsymbol{F}_{ij} + \sum_i \boldsymbol{F}'_i = \sum_i \boldsymbol{F}'_i.$$

En effet $\sum_{i,j} \boldsymbol{F}_{ij} = 0$, puisque relativement aux forces d'interaction on a $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$.

Corollaire 1. *La quantité de mouvement d'un système fermé est invariante.*

Corollaire 2. *Si la résultante des forces extérieures agissant sur un système est perpendiculaire à une droite x, la projection P_x de la quantité de mouvement sur la droite x est invariante*: $P_x =$ const.

Définition. On appelle *centre d'inertie* d'un système le point

$$\boldsymbol{r} = \frac{\sum m_i \boldsymbol{r}_i}{\sum m_i}.$$

Exercice. Montrer que le centre d'inertie est défini intrinsèquement, i. e. ne dépend pas du choix de l'origine des rayons vecteurs.

La quantité de mouvement d'un système est égale à la quantité de mouvement d'un point de masse Σm_i placé au centre d'inertie.

En effet, $(\sum m_i)\boldsymbol{r} = \sum (m_i \boldsymbol{r}_i)$, d'où $(\sum m_i)\dot{\boldsymbol{r}} = \sum m_i \dot{\boldsymbol{r}}_i$.

Désormais nous pouvons énoncer le théorème de la quantité de mouvement comme le théorème du mouvement du centre d'inertie.

Théorème. *Le centre d'inertie d'un système se déplace comme si toutes les masses étaient concentrées et toutes les forces appliquées en lui.*

Démonstration. $\left(\sum m_i\right)\dot{\boldsymbol{r}} = \boldsymbol{P}$, donc $\left(\sum m_i\right)\ddot{\boldsymbol{r}} = \frac{d\boldsymbol{P}}{dt} = \sum_i \boldsymbol{F}_i$.

Corollaire. *Si un système est fermé, son centre d'inertie est en mouvement rectiligne uniforme.*

C. Loi de conservation du moment cinétique.

Définition. On appelle *moment cinétique d'un point matériel par rapport à O* le moment du vecteur impulsion par rapport à O:

$$\boldsymbol{M} = [\boldsymbol{r}, m\dot{\boldsymbol{r}}].$$

Le *moment cinétique d'un système par rapport à O* est la somme des moments cinétiques de tous ses points:

$$\boldsymbol{M} = \sum_{i=1}^{n} [\boldsymbol{r}_i, m_i \dot{\boldsymbol{r}}_i].$$

Théorème. *La vitesse de variation du moment cinétique d'un système est égale à la somme des moments des forces extérieures agissant sur ce système.*

D é m o n s t r a t i o n.

$$\frac{dM}{dt}=\sum_{i=1}^{n}[\dot{r}_i,\ m_i\dot{r}_i]+\sum_{i=1}^{n}[r_i,\ m_i\ddot{r}_i].$$

Le premier terme est nul, le second est égal, en vertu des équations de Newton, à

$$\sum_{i=1}^{n}[r_i,\ F_i]=\sum_{i=1}^{n}[r_i,\ (\sum_{i\neq j}F_{ij}+F'_i)]=\sum_{i=1}^{n}[r_i,\ F'_i].$$

La somme des moments de deux forces d'interaction est nulle, puisque

$$F_{ij}=-F_{ji},\ [r_i,\ F_{ij}]+[r_j,\ F_{ji}]=[(r_i-r_j),\ F_{ij}]=0.$$

Donc est également nulle la somme des moments de toutes les forces d'interaction :

$$\sum_{i=1}^{n}\left[r_i,\ \sum_{i\neq j}F_{ij}\right]=0.\ \text{Finalement}\ \frac{dM}{dt}=\sum_{i=1}^{n}[r_i,\ F'_i],\ \text{c.q.f.d.}$$

Corollaire 1. (Loi de conservation du moment cinétique). *Si un système est fermé, alors* M = const.

Désignons la somme des moments des forces extérieures par $N=\sum_{i=1}^{n}[r_i,\ F'_i]$. Alors en vertu du théorème démontré $\frac{dM}{dt}=N$ et il résulte le

Corollaire 2. *Si le moment des forces extérieures par rapport à la droite z est nul, alors M_z est invariant.*

D. Loi de conservation de l'énergie.

D é f i n i t i o n. On appelle *énergie cinétique d'un point de masse m* l'expression

$$T=\frac{m\dot{r}^2}{2}.$$

D é f i n i t i o n. On appelle *énergie cinétique d'un système de points matériels* la somme des énergies cinétiques de chaque point :

$$T=\sum_{i=1}^{n}\frac{m_i\dot{r}_i^2}{2},$$

où m_i sont les masses des points et $\dot{r}_i$ leurs vitesses.

Théorème. *L'accroissement de l'énergie cinétique d'un système est égal à la somme des travaux de toutes les forces agissant sur ses points.*

D é m o n s t r a t i o n.

$$\frac{dT}{dt}=\sum_{i=1}^{n} m_i(\dot{\boldsymbol{r}}_i, \ddot{\boldsymbol{r}}_i)=\sum_{i=1}^{n}(\dot{\boldsymbol{r}}_i, m_i\ddot{\boldsymbol{r}}_i)=\sum_{i=1}^{n}(\dot{\boldsymbol{r}}_i, \boldsymbol{F}_i),$$

donc

$$T(t)-T(t_0)=\int_{t_0}^{t}\frac{dT}{dt}\,dt=\sum_{i=1}^{n}\int_{t_0}^{t}(\dot{\boldsymbol{r}}_i, \boldsymbol{F}_i)\,dt=\sum_{i=1}^{n}A_i, \text{ c.q.f.d.}$$

L'espace de configuration d'un système de n points matériels dans E^3 est le produit direct de n espaces euclidiens : $E^{3n} = E^3 \times \ldots \times E^3$. Il possède une structure d'espace euclidien. Soit $\boldsymbol{r} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$ le rayon vecteur d'un point de l'espace de configuration et $\boldsymbol{F} = (\boldsymbol{F}_1, \ldots, \boldsymbol{F}_n)$ un vecteur force. Le théorème précédent peut s'écrire sous la forme

$$T(t_1)-T(t_0)=\int_{\boldsymbol{r}(t_0)}^{\boldsymbol{r}(t_1)}(\boldsymbol{F}, d\boldsymbol{r})=\int_{t_0}^{t_1}(\dot{\boldsymbol{r}}, \boldsymbol{F})\,dt.$$

En d'autres termes:

L'accroissement de l'énergie cinétique est égal au travail de la « force » $\boldsymbol{F}$ de dimension $3n$ pour un « chemin » $\boldsymbol{r}(t)$ dans l'espace de configuration.

D é f i n i t i o n. On dit qu'un système est *potentiel* (ou *conservatif*) si les forces dépendent uniquement de la position de ses points: $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{r})$ et le travail de $\boldsymbol{F}$ pour un chemin quelconque dépend uniquement des extrémités de ce chemin:

$$\int_{M_1}^{M_2}(\boldsymbol{F}, d\boldsymbol{r})=\Phi(M_1, M_2).$$

Théorème. *Une condition nécessaire et suffisante de potentialité d'un système est l'existence de l'énergie potentielle, i. e. d'une fonction $U(\boldsymbol{r})$ telle que*

$$\boldsymbol{F}=-\frac{\partial U}{\partial \boldsymbol{r}}.$$

D é m o n s t r a t i o n. Voir le § 4, B.

Théorème. *L'énergie totale $E = T + U$ d'un système potentiel se conserve pendant le mouvement: $E(t_1) = E(t_0)$.*

D é m o n s t r a t i o n. On a démontré plus haut

$$T(t_1)-T(t_0)=\int_{\boldsymbol{r}(t_0)}^{\boldsymbol{r}(t_1)}(\boldsymbol{F}, d\boldsymbol{r})=U(\boldsymbol{r}(t_0))-U(\boldsymbol{r}(t_1)), \text{ c.q.f.d.}$$

Supposons que toutes les forces agissant sur les points du système se décomposent en forces d'interaction et en forces extérieures:

$$\boldsymbol{F}_i = \sum_{i \neq j} \boldsymbol{F}_{ij} + \boldsymbol{F}'_i,$$

où $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji} = f_{ij}\boldsymbol{r}_{ij}$.

A s s e r t i o n. *Si les forces d'interaction sont fonctions seulement des distances,* $f_{ij} = f_{ij}(|\boldsymbol{r}_i - \boldsymbol{r}_j|)$, *alors elles sont potentielles.*

D é m o n s t r a t i o n. Si le système est composé uniquement de deux points i, j, alors on vérifie aisément que l'énergie potentielle d'interaction est donnée par la formule

$$U_{ij}(r) = -\int_{r_0}^{r} f_{ij}(\rho)\, d\rho.$$

En effet, on aura

$$-\frac{\partial U_{ij}(|\boldsymbol{r}_i - \boldsymbol{r}_j|)}{\partial \boldsymbol{r}_i} = f_{ij}\frac{\partial |\boldsymbol{r}_i - \boldsymbol{r}_j|}{\partial \boldsymbol{r}_i} = f_{ij}\boldsymbol{e}_{ij}.$$

Donc l'énergie potentielle d'interaction de tous les points sera

$$U(r) = \sum_{i>j} U_{ij}(|\boldsymbol{r}_i - \boldsymbol{r}_j|), \text{ c.q.f.d.}$$

Si les forces extérieures sont potentielles, i.e. $\boldsymbol{F}'_i = -\frac{\partial U'_i}{\partial \boldsymbol{r}_i}$, alors le système est potentiel et son énergie potentielle totale est

$$U(r) = \sum_{i>j} U_{ij} + \sum_i U'_i.$$

L'énergie mécanique totale d'un tel système aura pour expression

$$E = T + U = \sum_i \frac{\dot{\boldsymbol{r}}_i^2}{2} + \sum_{i>j} U_{ij} + \sum_i U'_i.$$

Si le système n'est pas potentiel, alors son énergie mécanique totale d'une façon générale ne se conserve pas.

D é f i n i t i o n. Le décroissement de l'énergie mécanique $E(t_0) - E(t_1)$ s'appelle accroissement de l'*énergie non mécanique* E':

$$E'(t_1) - E'(t_0) = E(t_0) - E(t_1).$$

Théorème (loi de conservation de l'énergie). *L'énergie totale* $H = E + E'$ *se conserve.*

Il est bien entendu que ce théorème est une conséquence immédiate de la définition précédente. L'intérêt qu'il présente est que dans des systèmes physiques concrets l'énergie non mécanique E' est exprimée en fonction d'autres grandeurs physiques (température, etc.).

E. Exemple. Problème des deux corps. Soient deux points de masses m_1 et m_2 se déplaçant dans un champ de potentiel U; leurs équations du mouvement

$$m_1\ddot{r}_1 = -\frac{\partial U}{\partial r_1}, \quad m_2\ddot{r}_2 = -\frac{\partial U}{\partial r_2}, \quad U = U(|r_1 - r_2|).$$

Théorème. *La variation $r = r_1 - r_2$ dans le problème des deux corps est la même que dans le mouvement d'un point de masse $m = \frac{m_1 m_2}{m_1+m_2}$ dans un champ de potentiel $U(|r|)$.*

Soit r_0 le rayon vecteur du centre d'inertie

$$r_0 = \frac{m_1 r_1 + m_2 r_2}{m_1 + m_2}.$$

En vertu du théorème de conservation de la quantité de mouvement le point r_0 est en mouvement rectiligne uniforme.

Considérons maintenant le vecteur $r = r_1 - r_2$. En multipliant la première des équations du mouvement par m_2 et la seconde par m_1

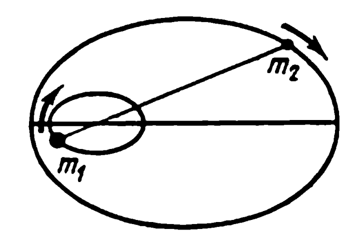


Fig. 40. Problèmes des deux corps.

et en soustrayant on obtient $m_1 m_2 \ddot{r} = -(m_1 + m_2)\frac{\partial U}{\partial r}$, où $U = U(|r_1 - r_2|) = U(|r|)$.

En particulier, dans le cas de l'attraction newtonienne, les points décrivent autour de leur centre d'inertie commun des coniques dont l'un des foyers est occupé par le centre d'inertie (fig. 40).

E x e r c i c e. Déterminer le demi grand axe de l'ellipse décrite par le centre de la Terre autour du centre d'inertie commun de la Terre et de la Lune. Où se trouve ce centre d'inertie: à l'intérieur ou à l'extérieur de la Terre? (La Lune est 81 fois moins dense que la Terre.)

§ 11. Similitude

Dans certains cas on peut recueillir une information importante sans résoudre les équations du mouvement, de leur seule forme, grâce à des critères de similitude et de dimension. Plus exactement, il s'agit de modifier l'échelle (du temps, de la longueur, de la masse, etc.) de telle façon que les équations du mouvement conservent leur forme.

A. Exemple. Supposons que $r(t)$ satisfait l'équation

$$m\frac{d^2r}{dt^2} = -\frac{\partial U}{\partial r}.$$

Posons $t_1 = \alpha t$, $m_1 = \alpha^2 m$. Alors $\boldsymbol{r}(t_1)$ vérifie également l'équation $m_1 \cdot \frac{d^2 r}{dt_1} = -\frac{\partial U}{\partial r}$. En d'autres termes:

Si l'on réduit la masse du point de quatre fois, il parcourra deux fois plus vite *) *la même orbite dans le même champ de forces.*

B. Exercice. Supposons que l'énergie potentielle d'un champ central est une fonction homogène de degré ν:

$$U(\alpha r) = \alpha^\nu U(r) \quad \forall \alpha > 0.$$

Montrer que *si une courbe γ est orbite d'un mouvement, la courbe homothétique $\alpha\gamma$ le sera également* (avec des conditions initiales dûment choisies). Trouver le rapport des temps de révolution sur ces orbites. En déduire que les oscillations du pendule sont isochrones ($\nu = 2$) et la troisième loi de Kepler ($\nu = -1$).

Exercice. Sachant que le rayon d'une planète est α fois plus petit que celui de la Terre et sa masse β fois plus petite, trouver de combien de fois l'accélération de la pesanteur, la première et la deuxième vitesse cosmique sont-elles plus petites que sur la Terre.

Réponse. $\gamma = \beta\alpha^{-2}$, $\delta = \sqrt{\beta/\alpha}$.

Par exemple, pour la Lune $\alpha \approx 3{,}7$, $\beta \approx 81$. Donc l'accélération de la pesanteur est environ le $^1/_6$ de celle de la Terre ($\gamma \approx 6$) et les vitesses cosmiques environ le $^1/_5$ de celles de la Terre ($\delta \approx 4{,}7$).

Exercice **). Les animaux vivant dans le désert doivent franchir de grandes distances entre les sources d'eau. Comment le temps maximal pendant lequel un animal est capable de courir dépend-il de sa taille L?

Réponse. Le temps est directement proportionnel à L.

Solution. Les réserves d'eau sont proportionnelles au volume du corps, i.e. à L^3; l'évaporation est proportionnelle à la surface, i.e. à L^2. Donc la durée maximale de course d'une source à l'autre est directement proportionnelle à L.

On remarquera que la distance maximale qu'est susceptible de parcourir un animal est également proportionnelle à L (voir l'exercice suivant).

Exercice ***). Comment varie la vitesse v d'un animal de taille L en terrain plat et sur une côte?

Réponse. En terrain plat $v \sim L^0$, sur une côte $v \sim L^{-1}$.

Solution. La puissance développée par un animal est proportionnelle à L^2 (le rendement des muscles est à peu près constant, environ 25 %, les 75 % restants de l'énergie chimique se transforment en chaleur; le dégagement de chaleur est proportionnel à la surface du corps, i.e. à L^2, donc la puissance utile est également proportionnelle à L^2).

La résistance de l'air est directement proportionnelle au carré de la vitesse et à l'aire de la section transversale; c'est pourquoi la puissance dépensée pour vaincre cette résistance est proportionnelle à v^2L^2v. Donc $v^3L^2 \sim L^2$ et $v \sim L^0$.

---

*) On suppose ici que U ne dépend pas de m. Dans le champ de gravitation l'énergie potentielle U est proportionnelle à m, et pour cette raison la période ne dépend pas de la masse m du point en mouvement.

**) J. Smith, *Mathematical Ideas in Biology*. Cambridge, 1968.

***) *Ibidem*.

En effet, en terrain plat la vitesse des animaux pas plus petits qu'un lièvre et pas plus grands qu'un cheval ne dépend pas de la taille du sujet.

Pour grimper une côte il faut développer une puissance $mgv \sim L^3 v$; comme la puissance développée est proportionnelle à L^2, il vient $v \sim L^{-1}$. Et effectivement le chien aborde avec facilité une côte, tandis que le cheval ralentit sa course.

E x e r c i c e *). Comment varie la hauteur du saut d'un animal en fonction de sa taille?

Réponse. $\sim L^0$.

S o l u t i o n. Pour effectuer un saut d'une hauteur h il faut une énergie potentielle proportionnelle à $L^3 h$; le travail de la force F développée par les muscles est proportionnel à FL. La force F est à son tour proportionnelle à L^2 (puisque la résistance des os est proportionnelle à leur section). Donc $L^3 h \sim L^2 L$, i.e. la hauteur du saut ne dépend pas de la taille de l'animal. Et effectivement la gerboise et le kangourou font des sauts de même hauteur.

---

*) J. S m i t h, *Mathematical Ideas in Biology*. Cambridge, 1968.

DEUXIÈME PARTIE

MÉCANIQUE DE LAGRANGE

La mécanique de Lagrange décrit le mouvement d'un système mécanique à l'aide de l'espace de configuration. L'espace de configuration d'un système mécanique possède une structure de variété différentiable. Sur une variété différentiable opère un groupe de difféomorphismes. Les principales notions et théorèmes de la mécanique lagrangienne (même s'ils sont formulés en termes de coordonnées locales) sont invariants par ce groupe *).

Un système mécanique lagrangien est défini par une variété (« espace de configuration ») et une fonction sur son fibré tangent (« fonction de Lagrange »).

Tout groupe à un paramètre de difféomorphismes de l'espace de configuration laissant invariante la fonction da Lagrange définit une loi de conservation (i. e. une intégrale première des équations du mouvement).

Le système potentiel newtonien est un cas particulier du lagrangien (l'espace de configuration est dans ce cas euclidien et la fonction de Lagrange égale à la différence de l'énergie cinétique et de l'énergie potentielle).

Le point de vue lagrangien permet d'étudier entièrement des problèmes importants de mécanique, relevant par exemple de la théorie des petites oscillations et de la dynamique du solide.

---

*) Voire par un groupe plus vaste de transformations qui opèrent également sur le temps.

CHAPITRE 3

PRINCIPE VARIATIONNEL

Dans ce chapitre on démontre que les mouvements d'un système potentiel newtonien sont les extrémales d'un principe variationnel appelé « principe de moindre action de Hamilton ».

Ce fait entraîne de nombreux corollaires importants, notamment un procédé d'écriture rapide des équations de mouvement en coordonnées curvilignes, ainsi que des conclusions qualitatives tel le théorème du retour au voisinage du point initial.

Dans ce chapitre l'espace est un espace arithmétique de dimension n. Un vecteur $\boldsymbol{x}$ d'un tel espace est une collection de nombres $(x_1, \ldots, x_n)$. Respectivement $\partial f/\partial \boldsymbol{x}$ désigne $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$ et $(\boldsymbol{a}, \boldsymbol{b}) = a_1b_1 + \ldots + a_nb_n$.

§ 12. Calcul des variations

Dans ce paragraphe sont exposés quelques préliminaires de calcul des variations dont on aura besoin par la suite.

L'objet du calcul des variations est la recherche des extremums de fonctions dont le domaine de définition est un espace de dimension infinie : l'espace des courbes. Ces fonctions sont appelées *fonctionnelles.*

Un exemple de fonctionnelle nous est fourni par la longueur d'une courbe dans le plan euclidien

$$\gamma = \{t, x : x(t) = x;\ t_0 \leqslant t \leqslant t_1\},$$

$$\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2}\, dt.$$

D'une façon générale, on appelle fonctionnelle toute application de l'espace des courbes dans l'axe numérique.

Soit une courbe $\gamma' = \{t, x : x = x(t) + h(t)\}$ « voisine » de la courbe γ. Désignons-la par $\gamma' = \gamma + h$. Soit l'accroissement de la fonctionnelle Φ : $\Phi(\gamma + h) - \Phi(\gamma)$ (fig. 41).

A. Variations.

D é f i n i t i o n. Une fonctionnelle Φ est différentiable *) si $\Phi(\gamma + h) - \Phi(\gamma) = F + R$, où F dépend linéairement de h (i. e. pour γ fixe $F(h_1 + h_2) = F(h_1) + F(h_2)$, $F(ch) = cF(h)$)

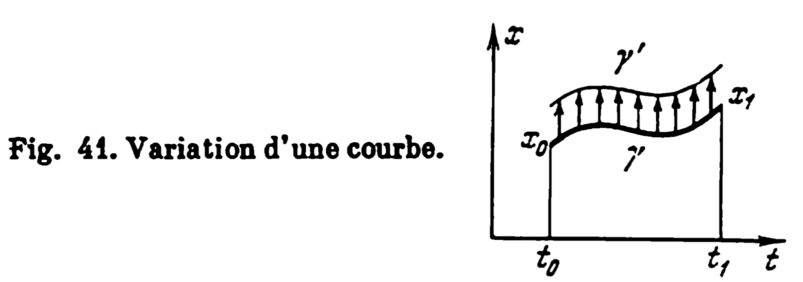


Fig. 41. Variation d'une courbe.

et $R(h, \gamma) = O(h^2)$ en ce sens que $|h| < \varepsilon$, $\left|\frac{dh}{dt}\right| < \varepsilon$, entraînent $|R| < C\varepsilon^2$. La partie linéaire $F(h)$ de l'accroissement s'appelle *différentielle.*

On démontre que si la fonctionnelle Φ est différentiable, alors sa différentielle est *univoquement* définie. Cette différentielle s'appelle aussi *variation* de la fonctionnelle et h, *variation de la courbe.*

E x e m p l e. Soit $\gamma = \{t, x : x = x(t),\ t_0 \leqslant t \leqslant t_1\}$ une courbe sur le plan t, x; $\dot{x} = \frac{dx}{dt}$; $L = L(a, b, c)$ une fonction différentiable par rapport à ses trois variables. Composons la fonctionnelle

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t)\, dt.$$

Dans le cas particulier $L = \sqrt{1 + b^2}$ on obtient la longueur de la courbe γ.

Théorème. *La fonctionnelle* $\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)\, dt$ *est différentiable et sa différentielle est donnée par la formule*

$$F(h) = \int_{t_0}^{t_1} \left[\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right] h\, dt + \left(\frac{\partial L}{\partial \dot{x}} h\right)\Big|_{t_0}^{t_1}.$$

*) En fait il aurait fallu indiquer sur quelle classe de courbes est définie la fonctionnelle Φ et quel espace vectoriel parcourt h. On peut supposer par exemple que dans les deux cas il s'agit de fonctions indéfiniment différentiables.

Démonstration.

$$\Phi(\gamma+h)-\Phi(\gamma)=\int_{t_0}^{t_1}[L(x+h,\dot{x}+\dot{h},t)-L(x,\dot{x},t)]\,dt=$$

$$=\int_{t_0}^{t_1}\left[\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}\right]dt+O(h^2)=F(h)+R,$$

où $F(h)=\int_{t_0}^{t_1}\left(\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}\right)dt$, $R=O(h^2)$. Une intégration par parties donne

$$\int_{t_0}^{t_1}\frac{\partial L}{\partial \dot{x}}\dot{h}\,dt=-\int_{t_0}^{t_1}h\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)dt+\left(h\frac{\partial L}{\partial \dot{x}}\right)\Big|_{t_0}^{t_1},\quad \text{c.q.f.d.}$$

B. Extrémales.

Définition. On appelle *extrémale* d'une fonctionnelle différentiable $\Phi(\gamma)$ une courbe γ telle que $F(h,\gamma)=0$ quel que soit h.

(Exactement comme γ est un point stationnaire d'une fonction si la différentielle de cette fonction s'annule en ce point.)

Théorème. *Une condition nécessaire et suffisante pour qu'une courbe* $\gamma: x=x(t)$ *soit extrémale de la fonctionnelle* $\Phi(\gamma)=$ $=\int_{t_0}^{t_1}L(x,\dot{x},t)\,dt$ *sur l'espace des courbes passant par les points* $x(t_0)=x_0$, $x(t_1)=x_1$ *est que sur la courbe* $x(t)$ *l'on ait*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)-\frac{\partial L}{\partial x}=0.$$

Lemme. *Si une fonction continue* $f(t)$, $t_0\leqslant t\leqslant t_1$, *est telle que* $\int_{t_0}^{t_1}f(t)h(t)\,dt=0$ *quelle que soit une fonction continue* *) $h(t)$ *telle que* $h(t_0)=h(t_1)=0$, *alors* $f(t)\equiv 0$.

Démonstration du lemme. Soit $f(t^*)>0$, $t_0<$ $<t^*<t_1$. En vertu de la continuité $f(t)>c$ dans un voisinage Δ

Fig. 42. Construction de la fonction h.

du point $t^*: t_0<t^*-d<t<t^*+d<t_1$; supposons que $h(t)=0$ à l'extérieur de Δ, $h(t)>0$ dans Δ et $h(t)=1$ dans

*) Ou au moins une fonction h indéfiniment différentiable.

$\frac{\Delta}{2}$ $\left(t^* - \frac{d}{2} < t < t^* + \frac{d}{2}\right)$. L'on a alors de toute évidence $\int_{t_0}^{t_1} f(t)\, h(t)\, dt \geqslant dc > 0$ (fig. 42). Cette contradiction montre que $f(t^*) = 0$ pour tous les $t_0 < t^* < t_1$, c.q.f.d.

D é m o n s t r a t i o n d u t h é o r è m e. En vertu du théorème précédent

$$F(h) = \int_{t_0}^{t_1} -\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x}\right] h\, dt + \left(\frac{\partial L}{\partial \dot{x}} h\right)\Big|_{t_0}^{t_1}.$$

Le terme hors de l'intégrale est nul, puisque $h(t_0) = h(t_1) = 0$. Si γ est extrémale, alors $F(h) = 0$ pour tous les h tels que $h(t_0) = h(t_1) = 0$. Donc pour tous ces $h(t)$ on a $\int_{t_0}^{t_1} f(t)\, h(t)\, dt = 0$, où $f(t) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x}$. D'après le lemme $f(t) \equiv 0$. Réciproquement si $f(t) \equiv 0$, il est évident que $F(h) \equiv 0$, c.q.f.d.

E x e m p l e. Vérifions que les extrémales de la longueur sont des droites. On a

$$L = \sqrt{1+\dot{x}^2}, \quad \frac{\partial L}{\partial x} = 0; \quad \frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1+\dot{x}^2}}, \quad \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{1+\dot{x}^2}}\right) = 0,$$

$$\frac{\dot{x}}{\sqrt{1+\dot{x}^2}} = c, \quad \dot{x} = c_1, \quad x = c_1 t + c_2.$$

C. Equation d'Euler-Lagrange.

D é f i n i t i o n. L'équation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0$$

est appelée *équation d'Euler-Lagrange pour la fonctionnelle*

$$\Phi = \int_{t_0}^{t_1} L(x, \dot{x}, t)\, dt.$$

Soit maintenant x un vecteur d'un espace arithmétique $\mathbf{R}^n$ de dimension n; $\gamma = \{t, x : x = x(t),\ t_0 \leqslant t \leqslant t_1\}$ une courbe de l'espace $\mathbf{R} \times \mathbf{R}^n$ de dimension $n + 1$ et $L: \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$ une fonction de $2n + 1$ arguments. Comme précédemment on démontre le

Théorème. *Une condition nécessaire et suffisante pour qu'une courbe* γ *soit extrémale de la fonctionnelle* $\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)\, dt$ *sur l'espace des courbes* $x(t)$ *passant par deux points donnés* (t_0, x_0) *et* (t_1, x_1) *est que l'équation d'Euler-Lagrange*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0$$

soit vérifiée sur cette courbe.

C'est un système de n équations du second ordre et la solution dépend de $2n$ constantes variables. Pour les trouver on dispose de $2n$ conditions $x(t_0) = x_0$, $x(t_1) = x_1$.

Exercice. Citer des exemples où il existe plusieurs extrémales passant par deux points et des exemples où il n'en existe pas.

D. Remarque importante. *La propriété d'une courbe* γ *d'être extrémale d'une fonctionnelle ne dépend pas du système de coordonnées.*

Exemple: une même fonctionnelle — la longueur d'une courbe — est donnée par des formules différentes en coordonnées cartésiennes et en coordonnées polaires:

$$\Phi_{\text{c.c.}} = \int_{t_0}^{t_1} \sqrt{\dot{x}_1^2 + \dot{x}_2^2}\, dt, \qquad \Phi_{\text{c.p.}} = \int_{t_0}^{t_1} \sqrt{\dot{r}^2 + r^2\dot{\varphi}^2}\, dt.$$

Les extrémales sont les mêmes: des droites dans le plan. Les équations des droites en coordonnées cartésiennes et en coordonnées polaires sont définies par des fonctions différentes: $x_1 = x_1(t)$, $x_2 = x_2(t)$; $r = r(t)$, $\varphi = \varphi(t)$. Cependant les unes comme les autres vérifient l'équation d'Euler-Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0$$

sauf que dans le premier cas $x = x_1, x_2$; $L = \sqrt{\dot{x}_1^2 + \dot{x}_2^2}$ et dans le second $x = r, \varphi$; $L = \sqrt{\dot{r}^2 + r^2\dot{\varphi}^2}$.

Donc nous pouvons facilement écrire l'équation différentielle de la famille de toutes les droites dans des coordonnées quelconques.

Exercice. Ecrire l'équation différentielle de la famille de toutes les droites d'un plan en coordonnées polaires.

§ 13. Equation de Lagrange

On indiquera ici un principe variationnel dont les extrémales sont solutions des équations newtoniennes du mouvement d'un système potentiel.

Comparons les équations de la dynamique de Newton

$$\frac{d}{dt}(m_i \dot{r}_i) + \frac{\partial U}{\partial r_i} = 0 \tag{1}$$

avec l'équation d'Euler-Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)-\frac{\partial L}{\partial x}=0.$$

A. Principe de moindre action de Hamilton.

Théorème. *Les mouvements du système mécanique* (1) *coïncident avec les extrémales de la fonctionnelle* $\Phi(\gamma)=\int_{t_0}^{t_1} L\,dt$, *où* $L = T - U$ *est la différence entre l'énergie cinétique et l'énergie potentielle.*

Démonstration. Comme $U = U(r)$ et $T = \sum m_i \frac{\dot{r}_i^2}{2}$, on a

$$\frac{\partial L}{\partial \dot{r}_i}=\frac{\partial T}{\partial \dot{r}_i}=m_i\dot{r}_i, \qquad \frac{\partial L}{\partial r_i}=-\frac{\partial U}{\partial r_i}.$$

Corollaire. *Soient* $(q_1, \ldots, q_{3n})$ *des coordonnées quelconques sur l'espace de configuration d'un système de n points matériels. Alors les variations de* q *en fonction du temps obéissent à l'équation d'Euler-Lagrange*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)-\frac{\partial L}{\partial q}=0, \quad \text{où } L = T - U.$$

Démonstration. D'après le théorème précédent, le mouvement est une extrémale de la fonctionnelle $\int L\,dt$. Donc dans tout système de coordonnées est vérifiée l'équation d'Euler-Lagrange transcrite dans ce système, c.q.f.d.

Définition. Les notations suivantes sont en usage en mécanique : $L(q, \dot{q}, t) = T - U$, *la fonction de Lagrange ou lagrangien* ; q_i, *les coordonnées généralisées*, $\dot{q}_i$, *les vitesses généralisées* ; $\frac{\partial L}{\partial \dot{q}_i} = p_i$, *les impulsions généralisées* ; $\frac{\partial L}{\partial q_i}$ *les forces généralisées* ; $\int_{t_0}^{t_1} L(q, \dot{q}, t)\,dt$, *l'action* ; $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)-\frac{\partial L}{\partial q_i}=0$, *les équations de Lagrange.*

Le dernier théorème s'appelle aussi « principe de moindre action de Hamilton », car quelquefois non seulement le mouvement $q(t)$ est extrémal, mais il minimise l'action $\int_{t_0}^{t_1} L\,dt$.

B. Exemples simples.

Exemple 1. Pour un point matériel libre dans E^3

$$L=T=\frac{m\dot{r}^2}{2};$$

en coordonnées cartésiennes $q_i=r_i$ et

$$L=\frac{m}{2}(\dot{q}_1^2+\dot{q}_2^2+\dot{q}_3^2).$$

Les vitesses généralisées sont les composantes du vecteur vitesse, les impulsions généralisées $p_i=m\dot{q}_i$ les composantes du vecteur de la quantité de mouvement ; les équations de Lagrange se confondent avec les équations de Newton $\frac{dp}{dt}=0$. Les extrémales sont des droites. Du principe de Hamilton il résulte que les *droites ne sont pas uniquement les plus courtes* (i.e. les extrémales de la longueur $\int_{t_0}^{t_1}\sqrt{\dot{q}_1^2+\dot{q}_2^2+\dot{q}_3^2}\,dt$), *elles sont également extrémales de l'action* $\int_{t_0}^{t_1}(\dot{q}_1^2+\dot{q}_2^2+\dot{q}_3^2)\,dt$.

Exercice. Démontrer que cet extremum est un minimum.

Exemple 2. Considérons le *mouvement* d'un point matériel libre *dans un champ central plan* en coordonnées polaires $q_1=r$, $q_2=\varphi$. De la relation $\dot{\boldsymbol{r}}=\dot{r}\boldsymbol{e}_r+\dot{\varphi}r\boldsymbol{e}_\varphi$ nous tirons l'énergie cinétique $T=\frac{m\dot{\boldsymbol{r}}^2}{2}=\frac{m}{2}(\dot{r}^2+r^2\dot{\varphi}^2)$ et le lagrangien $L(q,\dot{q})=T(q,\dot{q})-U(q)$, $U=U(q_1)$.

Les impulsions généralisées s'écrivent $p=\frac{\partial L}{\partial \dot{q}}$, i. e.

$$p_1=m\dot{r},\quad p_2=mr^2\dot{\varphi}.$$

La première équation de Lagrange $\dot{p}_1=\frac{\partial L}{\partial q_1}$ s'écrit

$$m\ddot{r}=mr\dot{\varphi}^2-\frac{\partial U}{\partial r}.$$

On reconnaît ici l'équation obtenue au § 8.

Comme $q_2=\varphi$ ne figure pas dans L, on a $\frac{\partial L}{\partial q_2}=0$. Et la deuxième équation de Lagrange s'écrira $\dot{p}_2=0$, $p_2=$ const. C'est la loi de conservation du moment cinétique.

Dans le cas général, lorsque le champ n'est pas central $U = U(r, \varphi)$ et $p_2 = -\frac{\partial U}{\partial \varphi}$.

Cette équation peut se mettre sous la forme $\frac{d}{dt}(\boldsymbol{M}, \boldsymbol{e}_z) = \boldsymbol{N}$, où $\boldsymbol{N} = ([\boldsymbol{r}, \boldsymbol{F}], \boldsymbol{e}_z)$, $\boldsymbol{F} = -\frac{\partial U}{\partial \boldsymbol{r}}$. (*La vitesse de variation du moment cinétique par rapport à la droite z est égale au moment de la force* $\boldsymbol{F}$ *par rapport à z.*)

En effet $dU = \frac{\partial U}{\partial r} dr + \frac{\partial U}{\partial \varphi} d\varphi = -(\boldsymbol{F}, d\boldsymbol{r}) = -(\boldsymbol{F}, \boldsymbol{e}_r)\, dr - r(\boldsymbol{F}, \boldsymbol{e}_\varphi)\, d\varphi$, donc $-\frac{\partial U}{\partial \varphi} = r(\boldsymbol{F}, \boldsymbol{e}_\varphi) = r([\boldsymbol{e}_r, \boldsymbol{F}], \boldsymbol{e}_z) = ([\boldsymbol{r}, \boldsymbol{F}], \boldsymbol{e}_z)$.

L'exemple traité suggère la généralisation suivante de la loi de conservation du moment cinétique.

Définition. La coordonnée q_i est appelée *cyclique* si elle ne figure pas dans la fonction de Lagrange: $\frac{\partial L}{\partial q_i} = 0$.

Théorème. *L'impulsion généralisée correspondant à une coordonnée cyclique se conserve*: $p_i = \text{const}$.

Démonstration. En vertu de l'équation de Lagrange

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} = 0, \text{ c.q.f.d:}$$

§ 14. Transformation de Legendre

La transformation de Legendre est un procédé mathématique auxiliaire consistant à passer des fonctions définies sur un espace vectoriel à des fonctions sur l'espace dual. La transformation de Legendre rappelle la dualité projective et les coordonnées tangentielles en géométrie algébrique ou encore la construction d'un espace de Banach adjoint en analyse. On la rencontre souvent en physique (par exemple, dans la définition des grandeurs thermodynamiques).

A. Définition. Soit $y = f(x)$ une fonction convexe, $f''(x) > 0$. On appelle *transformée de Legendre* de la fonction f une fonction g

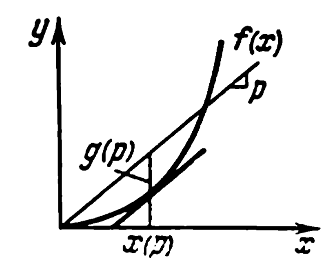


Fig. 43. Transformation de Legendre.

d'une nouvelle variable p construite de la façon suivante (fig. 43). Traçons sur le plan x, y le graphe de la fonction f. Soit donné un nombre p. Considérons la droite $y = px$. Prenons le point $x = x(p)$,

où la courbe est la plus éloignée de la droite $y = px$ à la verticale: la fonction $px - f(x) = F(p, x)$ admet au point $x(p)$ un maximum en x pour p fixe; on pose $g(p) = F(p, x(p))$.

Le point $x(p)$ est défini à partir de la condition d'extremum: $\partial F/\partial x = 0$, i. e. $f'(x) = p$. Le point $x(p)$ est unique *) en vertu de la convexité de f.

B. Exemples.

Exemple 1. Soit $f(x)=x^2$. Alors $F(p, x)=px-x^2$, $x(p)=\frac{p}{2}$, $g(p)=\frac{1}{4}p^2$.

Exemple 2. Soit $f(x)=\frac{mx^2}{2}$. Alors $g(p)=\frac{p^2}{2m}$.

Exemple 3. Soit $f(x)=\frac{x^\alpha}{\alpha}$. Alors $g(p)=\frac{p^\beta}{\beta}$, où $\frac{1}{\alpha}+\frac{1}{\beta}=1$ $(\alpha>1,\ \beta>1)$.

Exemple 4. Soit $f(x)$ une ligne polygonale convexe. Alors $g(p)$ est également une ligne polygonale convexe telle qu'aux sommets de $f(x)$ correspondent les segments de $g(p)$ et aux segments de $f(x)$ les sommets de $g(p)$. Par exemple à l'angle de la figure 44 la transformation de Legendre associe un segment.

C. Involution. On suppose que la fonction f est différentiable autant de fois qu'il le faut et $f''(x) > 0$. On vérifie sans peine que

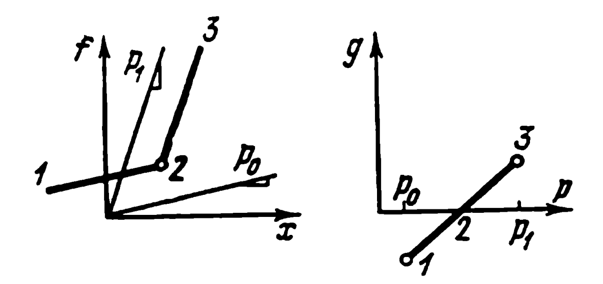


Fig. 44. La transformation de Legendre envoie un angle dans un segment.

la transformation de Legendre associe une fonction convexe à une fonction convexe. En l'appliquant deux fois on déduit le

Théorème. *La transformation de Legendre est involutive, i. e. son carré est égal à la transformation identique: si à f la transformation de Legendre associe g, alors à g elle associe de nouveau f.*

*) S'il existe.

Exercice. Montrer que le domaine de définition de g peut être constitué d'un point, d'un segment ou d'une demi-droite si la fonction f est définie sur l'axe x tout entier. Montrer que si la fonction f est définie sur un segment, alors la fonction g est définie sur l'axe p tout entier

D é m o n s t r a t i o n. Pour construire la transformée de Legendre de la fonction $g(p)$, nous devons par définition considérer une nouvelle variable indépendante (désignons-la par x), composer la fonction

$$G(x, p) = xp - g(p),$$

trouver le point $p(x)$ en lequel G admet un maximum

$$\frac{\partial G}{\partial p} = 0, \quad \text{i.e.} \quad g'(p) = x,$$

et alors la transformée de Legendre de $g(p)$ est une fonction de x égale à $G(x, p(x))$.

Montrons que $G(x, p(x)) = f(x)$. Remarquons à cette fin que la fonction $G(x, p) = xp - g(p)$ admet une interprétation géométrique simple : c'est l'ordonnée de la tangente au graphe $f(x)$ de

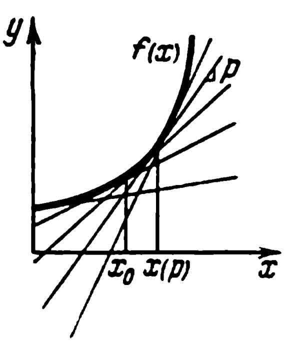


Fig. 45. Involutivité de la transformation de Legendre.

pente p, correspondant à l'abscisse x (fig. 45). En effet, pour p fixe, la fonction $G(x, p)$ est linéaire en x, en outre $\partial G/\partial x = p$ et pour $x = x(p)$ on a $G(x, p) = xp - g(p) = f(x)$ d'après la définition de $g(p)$.

Fixons maintenant $x = x_0$ et faisons varier p. Les valeurs de $G(x, p)$ seront alors les ordonnées des points d'intersection de la droite $x = x_0$ avec les tangentes au graphe $f(x)$ de pentes p diffés rentes. La convexité du graphe entraîne que *toutes ces tangente- sont situées sous la courbe*, et pour cette raison le maximum de $G(x, p)$ pour $x(p_0)$ fixe est égal à $f(x)$ (et est atteint pour $p = p(x_0) = f'(x_0)$), c.q.f.d.

Corollaire *). *Soit donnée une famille de droites $y = px - g(p)$. Alors l'enveloppe de cette famille a pour équation $y = f(x)$, où f est la transformée de Legendre de la fonction g.*

D. Inégalité de Young.

D é f i n i t i o n. Deux fonctions f et g qui se déduisent l'une de l'autre par la transformation de Legendre sont dites *duales au sens de Young*.

---

*) On reconnaît ici la théorie de l'« équation de Clairault ».

Par définition de la transformation de Legendre, $F(x, p) = px - f(x) \leqslant g(p)$ quels que soient x et p. De là résulte l'*inégalité de Young*

$$px \leqslant f(x) + g(p).$$

Exemple 1. Si $f(x) = \frac{x^2}{2}$, alors $g(p) = \frac{p^2}{2}$ et nous obtenons l'inégalité connue $px \leqslant \frac{x^2}{2} + \frac{p^2}{2}$ pour tous les x et p.

Exemple 2. Si $f(x) = \frac{x^\alpha}{\alpha}$, alors $g(p) = \frac{p^\beta}{\beta}$, $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ et nous obtenons l'inégalité de Young

$$px \leqslant \frac{x^\alpha}{\alpha} + \frac{p^\beta}{\beta}$$

pour tous les $x > 0$, $p > 0$, $\alpha > 1$, $\beta > 1$, $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

E. Cas de plusieurs variables. Soit maintenant $f(\boldsymbol{x})$ une fonction convexe de la variable vectorielle $\boldsymbol{x} = (x_1, \ldots, x_n)$ (i. e. la forme quadratique $\left(\frac{\partial^2 f}{\partial \boldsymbol{x}^2} d\boldsymbol{x}, d\boldsymbol{x}\right)$ est définie positive). On appelle transformée de Legendre la fonction $g(\boldsymbol{p})$ de la variable vectorielle $\boldsymbol{p} = (p_1, \ldots, p_n)$, définie comme précédemment par les égalités

$$g(\boldsymbol{p}) = F(\boldsymbol{p}, \boldsymbol{x}(\boldsymbol{p})) = \max_{\boldsymbol{x}} F(\boldsymbol{p}, \boldsymbol{x}), \quad F(\boldsymbol{p}, \boldsymbol{x}) = (\boldsymbol{p}, \boldsymbol{x}) - f(\boldsymbol{x}),$$

$$\boldsymbol{p} = \frac{\partial f}{\partial \boldsymbol{x}}.$$

Tous les raisonnements précédents y compris l'inégalité de Young se transposent intégralement à ce cas.

Exercice. Soit $f\colon \mathbf{R}^n \to \mathbf{R}$ une fonction convexe définie sur l'espace vectoriel $\mathbf{R}^n$. Désignons pa $\mathbf{R}^{n*}$ son espace dual. Montrer que les formules précédentes définissent une application $g\colon \mathbf{R}^{n*} \to \mathbf{R}$ (sous réserve que la forme linéaire $df|_{\boldsymbol{x}}$ parcourt l'espace $\mathbf{R}^{n*}$ tout entier lorsque $\boldsymbol{x}$ parcourt l'espace $\mathbf{R}^n$).

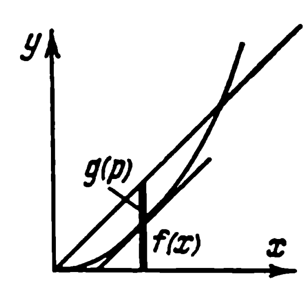


Fig. 46. Transformée de Legendre d'une forme quadratique.

E x e r c i c e. Soit f une forme quadratique : $f(x) = \sum f_{ij}x_ix_j$. Montrer que la transformation de Legendre lui fait correspondre une forme quadratique $g(p) = \sum g_{ij}p_ip_j$ telle que ces deux formes prennent les mêmes valeurs aux points correspondants (fig. 46) :

$$f(x, (p)) = g(p), \quad g(p(x)) = f(x).$$

§ 15. Equations de Hamilton

La transformation de Legendre associe au système lagrangien d'équations différentielles du second ordre un système remarquable symétrique de $2n$ équations du premier ordre, appelé *système d'équations de Hamilton* ou encore *d'équations canoniques*.

A. Equivalence des équations de Lagrange et de Hamilton. Soit le système d'équations de Lagrange $\dot{p} = \partial L/\partial q$, où $p = \partial L/\partial \dot{q}$, défini par la fonction de Lagrange $L: \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$ que nous supposons convexe *) relativement au deuxième argument $\dot{q}$.

Théorème. *Le système d'équations de Lagrange est équivalent à un système de $2n$ équations du premier ordre* ou équations de Hamilton :

$$\dot{p} = -\frac{\partial H}{\partial q},$$
$$\dot{q} = \frac{\partial H}{\partial p},$$

où $H(p, q, t) = p\dot{q} - L(q, \dot{q}, t)$ est la transformée de Legendre de la fonction de Lagrange considérée comme une fonction de $\dot{q}$.

D é m o n s t r a t i o n. Par définition la transformée de Legendre de $L(q, \dot{q}, t)$ par rapport à $\dot{q}$ est la fonction $H(p) = p\dot{q} - L(\dot{q})$, où $\dot{q}$ est exprimé en fonction de p d'après la formule $p = \partial L/\partial \dot{q}$ et qui dépend encore des paramètres q et t. *La fonction H est la fonction de Hamilton ou hamiltonien.*

La différentielle totale de la fonction de Hamilton

$$dH = \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial t}dt$$

est égale à la différentielle totale $p\dot{q} - L$ pour $p = \dfrac{\partial L}{\partial \dot{q}}$:

$$dH = \dot{q}\,dp - \frac{\partial L}{\partial q}dq - \frac{\partial L}{\partial t}dt.$$

---

*) Dans les applications cette fonction convexe sera en général une forme quadratique définie positive.

Les expressions de dH doivent se confondre. Donc

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q}, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

En tenant compte de l'équation de Lagrange $\dot{p} = \partial L/\partial q$ on obtient les équations de Hamilton.

Donc si $q(t)$ vérifie les équations de Lagrange, alors $(p(t), q(t))$ vérifie les équations de Hamilton. La réciproque se démontre de la même manière. Donc les systèmes de Lagrange et de Hamilton sont équivalents, c.q.f.d.

Remarque. Le théorème démontré s'applique à tous les problèmes variationnels et pas seulement aux équations lagrangiennes de la mécanique.

B. Fonction de Hamilton et énergie.

Exemple. Supposons que les équations sont tout de même mécaniques et la fonction de Lagrange de la forme habituelle $L = T - U$, où l'énergie cinétique T est une forme quadratique en $\dot{q}$:

$$T = \frac{1}{2}\sum a_{ij}\dot{q}_i\dot{q}_j, \quad a_{ij} = a_{ji}(q, t); \quad U = U(q).$$

Théorème. *Sous les hypothèses faites la fonction de Hamilton H est l'énergie totale $H = T + U$.*

La démonstration est basée sur le lemme relatif à la transformée de Legendre d'une forme quadratique.

Lemme. *La forme quadratique $f(x)$ et sa transformée de Legendre $g(p)$ prennent les mêmes valeurs aux points correspondants: $f(x) = g(p)$.*

Exemple. Pour la forme $f(x) = x^2$ c'est la propriété connue de la tengente à une parabole. Pour la forme $f(x) = \frac{mx^2}{2}$ on a $p = mx$ et $g(p) = \frac{p^2}{2m} = \frac{mx^2}{2} = f(x)$.

Démonstration du lemme. D'après le théorème d'Euler sur les fonctions homogènes

$$\frac{\partial f}{\partial x} x = 2f.$$

Donc

$$g(p(x)) = px - f(x) = \frac{\partial f}{\partial x} x - f = 2f(x) - f(x) = f(x), \text{ c.q.f.d.}$$

Démonstration du théorème. En raisonnant comme dans le lemme on trouve

$$H = p\dot{q} - L = 2T - (T - U) = T + U, \text{ c.q.f.d.}$$

E x e m p l e. Dans le cas d'un mouvement de dimension 1

$$\ddot{q} = -\frac{\partial U}{\partial q}.$$

Alors $T = \frac{1}{2}\dot{q}^2$, $U = U(q)$, $p = \dot{q}$, $H = \frac{p^2}{2} + U(q)$ et les équations de Hamilton s'écrivent

$$\begin{cases} \dot{q} = p, \\ \dot{p} = -\frac{\partial H}{\partial q}. \end{cases}$$

Cet exemple permet de se rappeler immédiatement laquelle des deux équations de Hamilton est affectée du signe moins.

Le théorème de l'équivalence des équations du mouvement d'un système hamiltonien entraîne d'importants corollaires. Par exemple, la loi de conservation de l'énergie acquiert une forme simple:

Corollaire 1. *L'égalité $dH/dt = \partial H/\partial t$ est vérifiée, et en particulier pour les systèmes dont la fonction de Hamilton ne dépend pas explicitement du temps ($\partial H/\partial t = 0$), est réalisée la loi de conservation des valeurs de la fonction de Hamilton:* $H(\mathbf{p}(t), \mathbf{q}(t)) =$ const.

D é m o n s t r a t i o n. Considérons les variations de H sur la trajectoire $H(\mathbf{p}(t), \mathbf{q}(t), t)$. Alors en vertu des équations de Hamilton

$$\frac{dH}{dt} = \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}, \text{ c.q.f.d.}$$

C. Coordonnées cycliques. En étudiant le champ central nous avons vu que l'introduction des coordonnées polaires ramenait le problème traité à un problème de dimension 1. Il s'avère que toute symétrie permettant de choisir un système de coordonnées $\mathbf{q}$, tel que la fonction de Hamilton n'y dépende pas de certaines coordonnées, permet de trouver quelques intégrales premières et de réduire le nombre de coordonnées de ce problème.

D é f i n i t i o n. Si la coordonnée q_1 ne figure pas dans la fonction de Hamilton $H(p_1, p_2, \ldots, p_n; q_1, \ldots, q_n; t)$, i. e. $\partial H/\partial q_1 = 0$, on dit qu'elle est *cyclique* (ce terme provient d'un cas particulier: la coordonnée angulaire φ dans un champ central). Il est évident que la coordonnée q_1 est cyclique si et seulement si elle ne figure pas dans la fonction de Lagrange ($\partial L/\partial q_1 = 0$). La forme hamiltonienne des équations du mouvement entraîne le

Corollaire 2. *Soit q_1 une coordonnée cyclique. Alors p_1 est intégrale première. Les autres coordonnées varient dans le temps comme dans un système de $n-1$ coordonnées indépendantes $q_2, \ldots, q_n$ dont la fonction de Hamilton*

$$H(p_2, \ldots, p_n, q_2, \ldots, q_n, t, c)$$

dépend du paramètre $c = p_1$.

Démonstration. Posons $\boldsymbol{p}' = (p_2, \ldots, p_n)$, $\boldsymbol{q}' = (q_2, \ldots, q_n)$. Le système de Hamilton s'écrit alors

$$\frac{d}{dt}\boldsymbol{q}' = \frac{\partial H}{\partial \boldsymbol{p}'}, \qquad \frac{d}{dt} q_1 = \frac{\partial H}{\partial p_1},$$
$$\frac{d}{dt}\boldsymbol{p}' = -\frac{\partial H}{\partial \boldsymbol{q}'}, \qquad \frac{d}{dt} p_1 = 0.$$

La dernière équation traduit le fait que $p_1 = \text{const}$. Donc dans le système d'équations en $\boldsymbol{p}'$ et $\boldsymbol{q}'$, la grandeur p_1 ne figure qu'en qualité de paramètre dans la fonction de Hamilton. Une fois ce système de $2n - 2$ équations résolu, l'équation en q_1 devient

$$\frac{d}{dt} q_1 = f(t), \quad \text{où} \quad f(t) = \frac{\partial}{\partial p_1} H(p_1, \boldsymbol{p}'(t), \boldsymbol{q}'(t), t),$$

et s'intègre sans difficultés.

La presque totalité des problèmes résolus en mécanique l'ont été à l'aide du corollaire 2.

Corollaire 3. *Tout système à deux degrés de liberté* ($n = 2$) *possédant une coordonnée cyclique est intégrable.* En effet, dans ce cas le système en p' et q' est de dimension 1 et s'intègre immédiatement à l'aide de l'intégrale $H(p', q') = c$.

§ 16. Théorème de Liouville

Le flot des équations de Hamilton conserve le volume. Il en résulte par exemple qu'un système de Hamilton ne peut être asymptotiquement stable.

Considérons le cas simple où la fonction de Hamilton ne dépend pas explicitement du temps: $H = H(\boldsymbol{p}, \boldsymbol{q})$.

A. Flot.

Définition. On appelle *espace des phases* un espace de dimension $2n$ muni des coordonnées $p_1, \ldots, p_n; q_1, \ldots, q_n$.

Exemple. Dans le cas $n = 1$ c'est le plan des phases du système $\ddot{x} = -\frac{\partial U}{\partial x}$ traité au § 4.

De même que dans ce cas simple, les seconds membres des équations de Hamilton définissent un champ de vecteurs: en chaque point $(\boldsymbol{p}, \boldsymbol{q})$ de l'espace des phases est appliqué le vecteur $(-\partial H/\partial \boldsymbol{q}, \partial H/\partial \boldsymbol{p})$ de dimension $2n$. Supposons que chaque solution des équations de Hamilton est prolongeable sur l'axe des temps tout entier *).

*) Par exemple, il suffit que les ensembles de niveau d'énergie de la fonction H soient compacts.

D é f i n i t i o n. On appelle *flot* un groupe à un paramètre de transformations de l'espace des phases

$$g^t : (\boldsymbol{p}(0),\ \boldsymbol{q}(0)) \to (\boldsymbol{p}(t),\ \boldsymbol{q}(t)),$$

où $\boldsymbol{p}(t)$, $\boldsymbol{q}(t)$ est solution du système d'équations de Hamilton (fig. 47).

E x e r c i c e. Montrer que g^t est un groupe.

B. Théorème de Liouville. 1) *Le flot conserve le volume: i.e. pour tout domaine D on a* (fig. 48)

$$\text{vol. } g^t D = \text{vol. } D.$$

Nous allons montrer une proposition plus générale 2) due également à Liouville.

Soit donné un système d'équations différentielles ordinaires $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \ldots, x_n)$ dont les solutions sont prolongeables

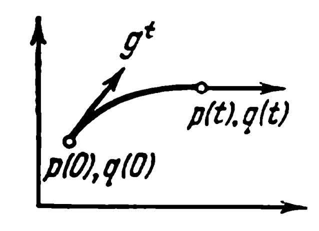


Fig. 47. Flot.

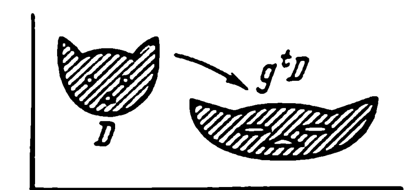


Fig. 48. Conservation du volume.

sur l'axe des temps tout entier. Soit g^t le groupe correspondant de transformations

$$g^t(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{f}(\boldsymbol{x})\,t + O(t^2) \quad (t \to 0). \tag{1}$$

Soit $D(0)$ un domaine de l'espace $\{\boldsymbol{x}\}$ et $v(0)$ son volume;

$$v(t) = \text{vol. } D(t), \quad D(t) = g^t D(0).$$

2) *Si* $\operatorname{div} \boldsymbol{f} \equiv 0$, *alors* g^t *conserve le volume*: $v(t) = v(0)$.

C. Démonstration.

Lemme 1. *On a la relation*

$$\left.\frac{dv}{dt}\right|_{t=0} = \int\limits_{D(0)} \operatorname{div} \boldsymbol{f}\, dx \quad (dx = dx_1 \ldots,\ dx_n).$$

D é m o n s t r a t i o n. D'après la définition du jacobien, pour t quelconque on a

$$v(t) = \int\limits_{D(0)} \det \frac{\partial g^t x}{\partial x}\, dx.$$

En calculant $\partial g^t x/\partial x$ moyennant la formule (1) on trouve pour $t \to 0$

$$\frac{\partial g^t x}{\partial x} = E + \frac{\partial f}{\partial x} t + O(t^2).$$

Utilisons maintenant un fait connu en algèbre.

Lemme 2. *Pour toute matrice $A = \| a_{ij} \|$ est vérifiée la relation*

$$\det | E + At | = 1 + t \operatorname{tr} A + O(t^2), \quad t \to 0,$$

où $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$ *est la trace de la matrice A (i. e. la somme des éléments diagonaux).*

(La démonstration du lemme 2 découle immédiatement du développement du déterminant: on obtient le nombre 1, n termes en t, les autres en t^2, t^3, etc.)

Donc

$$\det \frac{\partial g^t x}{\partial x} = 1 + t \operatorname{tr} \frac{\partial f}{\partial x} + O(t^2).$$

Or $\operatorname{tr} \frac{\partial f}{\partial x} = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} = \operatorname{div} f$, donc

$$v(t) = \int_{D(0)} [1 + t \operatorname{div} f + O(t^2)]\, dx,$$

ce qui démontre le lemme 1.

Démonstration du théorème 2). Comme $t = t_0$ n'est pas plus mauvais que $t = 0$, le lemme 1 peut s'écrire sous la forme

$$\left.\frac{dv(t)}{dt}\right|_{t=t_0} = \int_{D(t_0)} \operatorname{div} f \, dx.$$

Et si $\operatorname{div} f \equiv 0$, alors $\frac{dv}{dt} \equiv 0$, c.q.f.d.

En particulier, pour un système de Hamilton on a

$$\operatorname{div} f = \frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \frac{\partial}{\partial q}\left(\frac{\partial H}{\partial p}\right) \equiv 0.$$

Le théorème 1) de Liouville est démontré.

Exercice. Généraliser le théorème de Liouville au cas de systèmes non autonomes ($H = H(p, q, t)$ ou $f = f(x, t)$).

Exercice. Démontrer la formule de Liouville $W = W_0 \exp\left(\int \operatorname{tr} A \, dt\right)$ pour le wronskien du système linéaire $\dot{x} = A(t)\, x$.

Le théorème de Liouville possède d'innombrables applications.

Exercice. Montrer que ni la position d'équilibre d'un système hamiltonien ni son cycle limite ne peuvent être asymptotiquement stables dans l'espace des phases.

Les applications du théorème de Liouville sont les plus importantes en mécanique statistique.

Le théorème de Liouville permet d'appliquer aux systèmes mécaniques les méthodes de la *théorie ergodique* *). Voyons un exemple des plus simples.

D. Théorème du retour de Poincaré. *Soit g une application bijective continue conservant les volumes et associant à lui-même un domaine limité D d'un espace euclidien: $gD = D$.*

Alors dans tout voisinage U d'un point quelconque du domaine D il existe un point $x \in U$ qui retourne dans le domaine U, i. e. $g^n x \in U$ pour un certain $n > 0$.

Ce théorème est valable par exemple pour le flot g^t d'un système de dimension 2 dont le potentiel $U(x_1, x_2)$ tend vers l'infini; dans

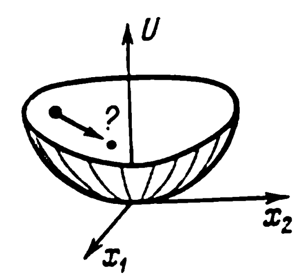


Fig. 49. Du mouvement de la bille dans la cuvette asymétrique la seule chose que l'on sache c'est que d'après le théorème de Poincaré la bille reviendra au voisinage de son point de départ.

ce cas un domaine invariant borné de l'espace des phases est défini par la condition (fig. 49):

$$D = \{p, q: T + U \leqslant E\}.$$

On peut renforcer le théorème de Poincaré en démontrant que tout point matériel en mouvement retourne plusieurs fois à sa position de départ. C'est l'une des rares conclusions générales que l'on puisse faire sur le caractère du mouvement. Même du mouvement simple

$$\ddot{x} = -\frac{\partial U}{\partial x}, \quad x = (x_1, x_2)$$

personne ne sait les détails.

Voyons un paradoxe découlant des théorèmes de Poincaré et de Liouville: si l'on ouvre la cloison séparant une chambre remplie de gaz et une chambre sous vide, on constate qu'au bout d'un moment les molécules de gaz se rassemblent de nouveau dans la première chambre (fig. 50).

*) Cf. par exemple: P. Halmos, *Lectures on Ergodic Theory*. Chicago, 1956.

L'explication de ce paradoxe est que ce « moment » est plus long que l'existence du système Solaire.

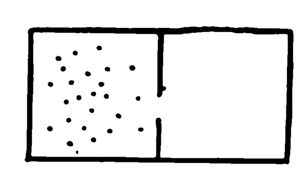


Fig. 50. Les molécules retournent dans la première chambre.

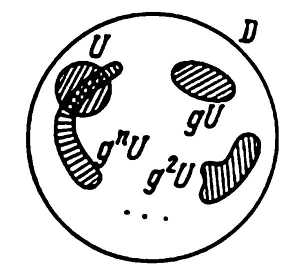


Fig. 51. Théorème du retour.

D é m o n s t r a t i o n d u t h é o r è m e d e P o i n c a r é. Soient les images du voisinage U (fig. 51):

$$U, gU, g^2U, \ldots, g^nU, \ldots$$

Ils possèdent tous un même volume positif. S'ils ne se coupaient pas, le volume D serait infini. Donc pour certains $k \geqslant 0$, $l \geqslant 0$, $k > l$

$$g^kU \cap g^lU \neq 0$$

et $g^{k-l}U \cap U \neq 0$. Soit $g^{k-l}x = y$, $x \in U$, $y \in U$. Alors $x \in U$, $g^nx \in U$ $(n = k - l)$, c.q.f.d.

E. Applications du théorème de Poincaré.

E x e m p l e 1. Soit D un cercle, g une rotation d'angle α. Si $\alpha = 2\pi \frac{m}{n}$, alors g^n est la transformation identique et le théorème est évident. Si α est

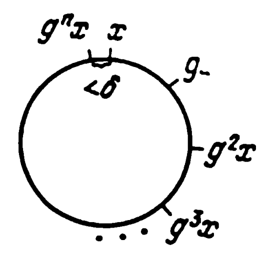


Fig. 52. Ensemble partout dense dans le cercle.

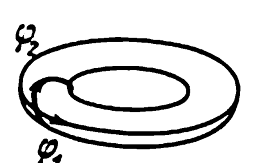


Fig. 53. Tore.

incommensurable avec 2π, alors le théorème de Poincaré donne

$$\forall\ \delta > 0, \quad \exists n: |\, g^n x - x \,| < \delta \text{ (fig. 52).}$$

D'où il découle immédiatement le

Théorème. *Si $\alpha \neq 2\pi \frac{m}{n}$, alors l'ensemble des points de la forme g^nx est partout dense *) dans un cercle* $(k = 1, 2, \ldots)$.

*) On rappelle qu'un ensemble A est partout dense dans B si tout voisinage d'un point de B contient au moins un point de A.

E x e r c i c e. Montrer que toute orbite d'un mouvement dans un champ central, avec $U = r^4$, soit est fermée, soit est partout dense dans une couronne.

E x e m p l e 2. Soit D un tore de dimension 2, sur lequel on définit des coordonnées angulaires φ_1 et φ_2 (latitude et longitude) (fig. 53).

Considérons le système d'équations différentielles ordinaires sur ce tore

$$\dot{\varphi}_1 = \alpha_1, \quad \dot{\varphi}_2 = \alpha_2.$$

De toute évidence $\operatorname{div} f = 0$ et le mouvement correspondant

$$g^t : (\varphi_1, \varphi_2) \to (\varphi_1 + \alpha_1 t, \varphi_2 + \alpha_2 t)$$

conserve le volume $d\varphi_1$, $d\varphi_2$. Le théorème de Poincaré entraîne aussitôt le

Théorème. *Si α_1/α_2 est irrationnel, alors l'hélice $g^t(\varphi_1, \varphi_2)$ est partout dense dans le tore D.*

E x e r c i c e. Montrer que si ω est irrationnel, alors la figure de Lissaous ($x = \cos t$, $y = \cos \omega t$) est partout dense dans le carré $|x| \leqslant 1$, $|y| \leqslant 1$.

E x e m p l e 3. Soit D un tore T^n de dimension n, i.e. le produit cartésien *) de n cercles:

$$D = \underbrace{S^1 \times S^1 \times \ldots \times S^1}_{n} = T^n.$$

Alors un point du tore T^n est donné par n coordonnées angulaires $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_n)$. Soit $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ et g^t une transformation conservant le volume

$$g^t : T^n \longrightarrow T^n, \quad \boldsymbol{\varphi} \longrightarrow \boldsymbol{\varphi} + \boldsymbol{\alpha} t.$$

E x e r c i c e. A quelles conditions doit satisfaire $\boldsymbol{\alpha}$ pour que soient partout denses dans T^n: a) la trajectoire $g^t\boldsymbol{\varphi}$, b) la trajectoire $g^k\boldsymbol{\varphi}$ ($t \in \mathbf{R}$ appartient à l'ensemble des nombres réels $\mathbf{R}$, $k \in \mathbf{Z}$ à l'ensemble des nombres entiers $\mathbf{Z}$).

Les transformations des exemples 1 à 3 sont étroitement liées à la mécanique. Or comme le théorème de Poincaré est abstrait, il possède également des applications dans des domaines autres que la mécanique.

E x e m p l e 4. Considérons les premiers chiffres des nombres 2^n: 1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4,...

E x e r c i c e. Est-ce que cette suite renferme le chiffre 7? Lequel des chiffres 7 et 8 possède la plus grande fréquence? Etablir le rapport de ces fréquences.

---

*) On appelle produit cartésien (ou direct) des ensembles A, B, ... l'ensemble des collections de points $(a, b, \ldots)$, $a \in A$, $b \in B$, ...

CHAPITRE 4

LA MÉCANIQUE DE LAGRANGE SUR LES VARIÉTÉS

Dans ce chapitre on introduit les notions de variété différentiable et de fibré tangent. La donnée de la fonction de Lagrange ou lagrangien sur le fibré tangent définit un « système holonome ». Les cas particuliers sont les systèmes de points matériels soumis à des liaisons holonomes, tels le pendule ou les solides.

§ 17. Liaisons holonomes

Dans ce paragraphe on définit ce qu'on entend par système de points matériels soumis à des liaisons holonomes.

A. Exemple. Soit γ une courbe différentiable sur le plan. Si au voisinage de γ agit un champ de forces très intense dirigé vers la courbe, alors ce champ maintiendra tout point en mouvement dans

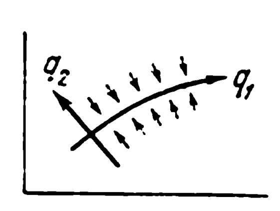


Fig. 54. La liaison comme un champ de forces infiniment intense.

Fig. 55. Energie potentielle U_N.

le voisinage de γ. Dans le cas limite d'un champ de forces d'une intensité infinie le point est astreint à se déplacer suivant la courbe γ. On dit alors que le système est soumis à une liaison (fig. 54).

Pour donner une formulation exacte considérons au voisinage de la courbe γ les coordonnées curvilignes q_1, q_2 : q_1 est portée par la courbe γ, q_2 est la distance à la courbe.

Soit un système dont l'énergie potentielle

$$U_N = Nq_2^2 + U_0(q_1, q_2)$$

dépend d'un paramètre N (qui tendra ensuite vers l'infini) (fig. 55).

Considérons les conditions initiales sur γ:

$$q_1(0) = q_1^0,\ \dot{q}_1(0) = \dot{q}_1^0,\ q_2(0) = 0,\ \dot{q}_2(0) = 0.$$

Désignons par $q_1 = \varphi(t, N)$ les variations de la coordonnée q_1 lorsque le point se déplace dans un mouvement satisfaisant aux conditions initiales ci-dessus dans le champ U_N.

Théorème. *Lorsque $N \to \infty$ existe la limite*

$$\lim_{N\to\infty} \varphi(t, N) = \psi(t).$$

La fonction limite $q_1 = \psi(t)$ vérifie l'équation de Lagrange

$$\frac{d}{dt}\left(\frac{\partial L_*}{\partial \dot{q}_1}\right) = \frac{\partial L_*}{\partial q_1},$$

où $L_*(q_1, \dot{q}_1) = T\,|_{q_2=\dot{q}_2=0} - U_0\,|_{q_2=0}$ (T est l'énergie cinétique du mouvement sur γ).

Donc lorsque $N \to \infty$, le système d'équations de Lagrange en q_1 et q_2 engendre une équation de Lagrange en $q_1 = \psi(t)$.

On obtiendrait exactement le même résultat si au lieu du plan on considérait l'espace de configuration de dimension $3n$ d'un système mécanique de n points, muni de la métrique $ds^2 = \sum_{i=1}^{n} m_i d\boldsymbol{r}_i^2$ (m_i sont les masses des points); au lieu de la courbe γ, une sous-variété de l'espace de configuration; au lieu de q_1, des coordonnées quelconques q_1 sur γ; au lieu de q_2, des coordonnées q_2 perpendiculaires à γ. Si l'énergie potentielle est de la forme

$$U = U_0(q_1) + Nq_2^2,$$

alors lorsque $N \to \infty$ le mouvement sur γ est défini par des équations de Lagrange dont le lagrangien est

$$L_* = T\,|_{q_2=\dot{q}_2=0} - U_0\,|_{q_2=0}.$$

B. Définition d'un système comportant des liaisons. Nous n'allons pas démontrer le théorème *) que nous venons d'énoncer et nous n'allons pas l'utiliser. Il nous servira à justifier la

D é f i n i t i o n. Soit γ une surface de dimension m dans l'espace de configuration de dimension $3n$ des points $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$ de masses respectives $m_1, \ldots, m_n$. Soit $\boldsymbol{q} = (q_1, \ldots, q_m)$ des coordonnées quelconques sur la courbe γ: $\boldsymbol{r}_i = \boldsymbol{r}_i(\boldsymbol{q})$. Le système décrit par les

*) La démonstration est basée sur le fait que, l'énergie se conservant, le point en mouvement ne peut s'éloigner de γ d'une distance supérieure à $cN^{-1/2}$, or cette quantité tend vers zéro lorsque $N \to \infty$.

équations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}=\frac{\partial L}{\partial q}, \quad L=\frac{1}{2}\sum m_i\dot{r}_i^2+U(q)$$

est appelé système de n points comportant $3n - m$ *liaisons holonomes* parfaites. La surface γ est l'*espace de configuration du système lié.*

Si la surface γ est donnée par $k = 3n - m$ équations fonctionnellement indépendantes $f_1(r) = 0, \ldots, f_k(r) = 0$, on dit que le système comporte les liaisons $f_1 = 0, \ldots, f_k = 0$.

Une liaison holonome pourrait se définir comme le cas limite d'un système doué d'une grande énergie potentielle. L'importance de ces liaisons réside dans le fait que, comme le montre l'expérience, de nombreux systèmes mécaniques appartiennent à cette classe avec une plus ou moins grande approximation.

Dans la suite nous appellerons simplement liaisons des liaisons holonomes parfaites. Ce seront les seules liaisons que nous aurons à considérer.

§ 18. Variétés différentiables

L'espace de configuration d'un système lié est une variété différentiable. Dans ce paragraphe on donne quelques généralités sur les variétés différentiables.

A. Définition d'une variété différentiable. Sur un ensemble M est définie une structure de variété différentiable si M est muni d'une collection finie ou dénombrable de *cartes* telles que chaque point soit représenté au moins sur une carte.

Par carte on désigne un domaine ouvert U d'un espace arithmétique euclidien $q = (q_1, \ldots, q_n)$ et une application bijective φ de U sur un sous-ensemble de M, $\varphi\colon U \to \varphi U \subset M$.

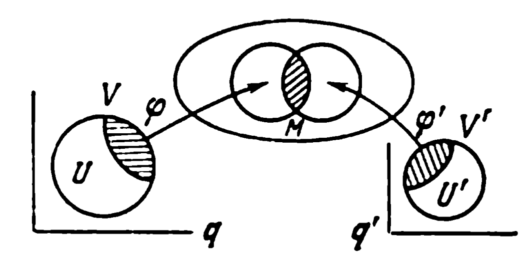


Fig. 56. Cartes compatibles.

Si un point de M est représenté sur deux cartes U et U' à la fois, il doit en être de même pour certains de ses voisinages V et V' sur chacune de ces cartes (fig. 56). On a donc l'application $\varphi'^{-1}\varphi$: $V \to V'$ d'une partie de la carte $V \subset U$ sur une partie de la carte $V' \subset U'$.

Ceci est une application du domaine V de l'espace arithmétique euclidien q sur le domaine V' de l'espace arithmétique q' ; elle est définie par n fonctions de n variables $q' = q'(q)$, $(q = q(q'))$. Les

cartes U et U' sont dites *compatibles* si ces fonctions sont différentiables *).

On appelle *atlas* un ensemble de cartes compatibles. Deux atlas sont *équivalents* si leur union est un atlas.

Une variété différentiable est une classe d'équivalence d'atlas. Nous n'étudierons que les variétés *connexes* **). Le nombre n sera le même pour toutes les cartes: on l'appelle *dimension* de la variété.

On appelle *voisinage* d'un point d'une variété l'image par l'application $\varphi: U \to M$ d'un voisinage de l'image de ce point sur une carte U. Nous supposerons que deux points différents d'une variété possèdent des voisinages disjoints.

B. Exemples.

Exemple 1. L'espace euclidien $\mathbf{R}^n$ est une variété dont l'atlas est composé d'une seule carte.

Exemple 2. La sphère $S^2 = \{(x, y, z): x^2 + y^2 + z^2 = 1\}$ possède une structure de variété dont l'atlas est composé par exemple de deux cartes

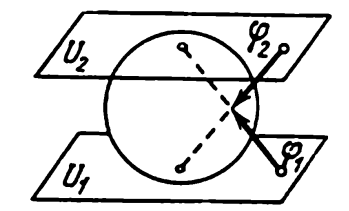


Fig. 57. Atlas de la sphère.

$(U_i, \varphi_i, i = 1, 2)$ en projection stéréographique (fig. 57). La construction est analogue dans le cas de la sphère de dimension n:

$$S^n = \{(x_1, \ldots, x_{n+1}): \sum x_i^2 = 1\}.$$

Exemple 3. Considérons un pendule plan. Son espace de configuration, un cercle S^1, est une variété. Un atlas est défini par la coordonnée angulaire $\varphi: \mathbf{R}^1 \to S^1$, $U_1 = (-\pi, \pi)$, $U_2 = (0, 2\pi)$ (fig. 58).

Exemple 4. L'espace de configuration du pendule mathématique sphérique est une sphère S^2 de dimension 2 (fig. 58).

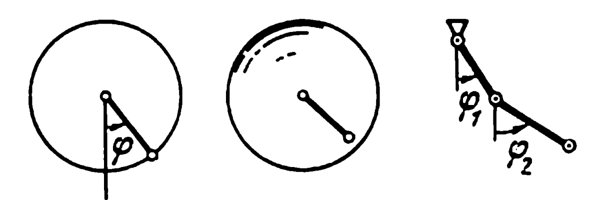


Fig. 58. Pendules plan, sphérique et double.

Exemple 5. L'espace de configuration du pendule double plan est le produit cartésien de deux cercles, i.e. un tore de dimension 2, $T^2 = S^1 \times S^1$ (fig. 58).

*) Par différentiabilité on entend ici une différentiabilité continue d'ordre r; la valeur exacte de r ($1 \leqslant r \leqslant \infty$) importe peu (on peut supposer que $r = \infty$, par exemple).

**) Une variété est connexe s'il est impossible de la décomposer en deux variétés disjointes.

E x e m p l e 6. L'espace de configuration du pendule double sphérique est le produit cartésien de deux sphères $S^2 \times S^2$.

E x e m p l e 7. Un segment rigide du plan q_1, q_2 a pour espace de configuration la variété $\mathbf{R}^2 \times S^1$ munie de coordonnées q_1, q_2, q_3 (fig. 59).

Elle est recouverte par deux cartes.

E x e m p l e 8. On considère un triangle rectangle rigide OAB en rotation autour de son sommet O. La position du triangle est donnée par trois nombres. En effet, la direction $OA \in S^2$ est définie par deux nombres, et si OA est donné, on peut faire tourner encore $OB \in S^1$ autour de l'axe OA (fig. 60).

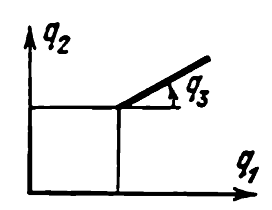


Fig. 59. Espace de configuration d'un segment sur le plan.

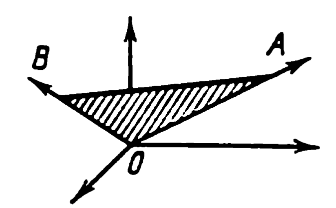


Fig. 60. Espace de configuration du triangle.

A chaque position du rectangle OAB peut être associé le repère orthonormé direct $e_1 = \frac{OA}{|OA|}$, $e_2 = \frac{OB}{|OB|}$, $e_3 = [e_1, e_2]$. Cette correspondance étant bijective, chaque position du triangle est définie par une matrice orthogonale d'ordre 3 et de déterminant 1.

L'ensemble de toutes les matrices d'ordre trois est un espace $\mathbf{R}^9$. Les six conditions d'orthogonalité définissent deux variétés connexes de dimension trois, de matrices de déterminant $+1$ et -1. Les rotations de l'espace ordinaire (de déterminant $+1$) forment un groupe noté SO (3).

Donc l'*espace de configuration du triangle OAB est le groupe* SO (3).

E x e r c i c e. Montrer que la variété SO (3) est homéomorphe à un espace projectif réel de dimension 3.

D é f i n i t i o n. La dimension de l'espace de configuration est appelée *nombre de degrés de liberté*.

E x e m p l e 9. Considérons un système de k tiges assemblées par articulation et formant une chaîne fermée.

E x e r c i c e. Quel est le nombre de degrés de liberté de ce système?

E x e m p l e 10. *Variété plongée*. On dit que M est une sous-variété de dimension k (fig. 61) plongée dans un espace euclidien E^n si

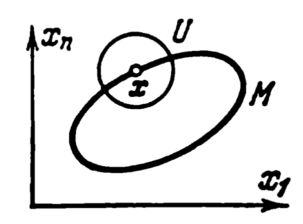


Fig. 61. Sous-variété plongée.

dans un voisinage U de chaque point $x \in M$ existent $n-k$ fonctions $f_1: U \to \mathbf{R}, \ldots, f_{n-k}: U \to \mathbf{R}$ telles que l'intersection du voisi-

nage U avec M soit définie par les équations $f_1 = 0, \ldots, f_{n-k} = 0$, et les vecteurs grad $f_1, \ldots$, grad f_{n-k} en x soient linéairement indépendants.

Il est facile de munir M d'une structure de variété, i.e. d'introduire des coordonnées au voisinage de x (comment ?).

On démontre que toute variété peut être plongée dans un espace euclidien.

Dans l'exemple 8 le groupe SO (3) est un sous-ensemble de $\mathbf{R}^9$.

E x e r c i c e. Montrer que SO (3) est plongé dans $\mathbf{R}^9$, en déduire que SO (3) est une variété.

C. Espace tangent. Si M est une variété de dimension k plongée dans E^n, alors en chacun de ses points x elle admet un espace tangent TM_x de dimension k. Plus exactement, TM_x est l'orthocomplément

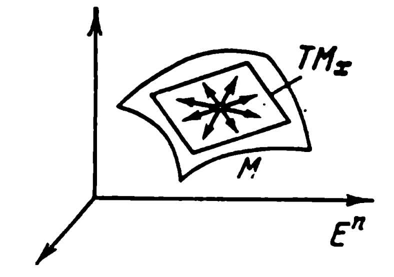


Fig. 62. Espace tangent.

de {grad $f_1, \ldots$, grad f_{n-k}} (fig. 62). Les vecteurs de l'espace tangent TM_x appliqués au point x sont appelés vecteurs tangents à M en x. Ces vecteurs peuvent être définis directement comme les vecteurs vitesses des courbes sur M:

$$\dot{x} = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t}, \quad \text{où} \quad \varphi(0) = x, \quad \varphi(t) \in M.$$

On pourrait définir un vecteur tangent en termes « internes », sans recourir au plongement de M dans E^n.

Nous dirons que deux courbes $x = \varphi(t)$ et $x = \psi(t)$ sur une variété sont *équivalentes* si $\varphi(0) = \psi(0) = x$ et $\lim\limits_{t \to 0} \frac{\varphi(t) - \psi(t)}{t} = 0$ sur une carte. Cette expression de la tangence vaut alors pour une carte quelconque (en faire la preuve !).

D é f i n i t i o n. On appelle *vecteur tangent* à la variété M en x la classe d'équivalence des courbes $\varphi(t)$, $\varphi(0) = x$.

On définit aisément le produit d'un vecteur tangent par un nombre et l'addition de vecteurs tangents. L'ensemble des vecteurs tangents en x à M formant un espace vectoriel TM_x est appelé *espace tangent en x à M*.

Pour les variétés plongées cette définition coïncide avec la précédente. Cependant elle a l'avantage de convenir également pour des variétés abstraites M nulle part plongées.

D é f i n i t i o n. Soit U une carte d'un atlas M muni des coordonnées $q_1, \ldots, q_n$. On appelle *composantes* d'un vecteur tangent à une courbe $q = \varphi(t)$ les nombres $\xi_1, \ldots, \xi_n$, où $\xi_i = \left.\frac{d\varphi_i}{dt}\right|_{t=0}$.

D. Fibré tangent. L'union des espaces tangents à une variété M en tous ses points, $\bigcup_{x \in M} TM_x$, possède une structure naturelle de variété différentiable dont la dimension est le double de celle de M.

Cette variété s'appelle *fibré tangent* à la variété M et se note TM. Un point de TM est un vecteur ξ tangent à M en un point x. Les coordonnées locales sur TM se construisent de la manière suivante. Soient $q_1, \ldots, q_n$ des coordonnées locales sur la variété M, $\xi_1, \ldots, \xi_n$ les composantes d'un vecteur tangent ξ dans ce système de coordonnées. Alors les $2n$ nombres $(q_1, \ldots, q_n, \xi_1, \ldots, \xi_n)$ définissent un système local de coordonnées sur TM.

L'application $p: TM \to M$ qui à chaque vecteur tangent ξ associe le point $x \in U$ en lequel ξ est tangent à M s'appelle *application projective*. L'antécédent du point $x \in M$ par l'application projective $p^{-1}(x)$ est l'espace tangent TM_x. Cet espace s'appelle *fibre du fibré au-dessus du point* x.

E. Variété riemannienne. Si M est une variété plongée dans un espace euclidien, alors la métrique de cet espace permet de mesurer sur M les longueurs des courbes, les angles des vecteurs, les volumes, etc.

Toutes ces grandeurs s'expriment en fonction du module des vecteurs tangents, i.e. en fonction d'une forme quadratique définie

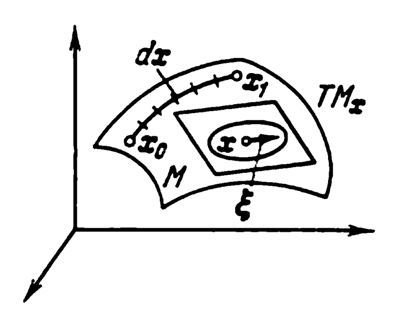


Fig. 63. Métrique riemannienne.

positive donnée sur chaque espace tangent TM_x (fig. 63):

$$TM_x \to \mathbf{R}; \quad \xi \to \langle \xi, \xi \rangle.$$

E x e m p l e: la longueur d'une courbe γ sur une variété s'exprime en fonction de cette forme comme $l(\gamma) = \int_{x_0}^{x_1} \sqrt{\langle dx, dx \rangle}$ ou si la courbe est définie paramétriquement :

$$\gamma : [t_0, t_1] \to M, \quad t \to x(t) \in M,$$

alors

$$l(\gamma)=\int_{t_0}^{t_1}\sqrt{\langle\dot{x},\dot{x}\rangle}\,dt.$$

D é f i n i t i o n. On appelle *variété riemannienne* une variété différentiable M de forme quadratique définie positive $\langle\xi,\xi\rangle$ fixe sur chaque espace tangent TM_x. Cette forme quadratique s'appelle *métrique riemannienne*.

R e m a r q u e. Soit U une carte d'un atlas de M muni des coordonnées $q_1, \ldots, q_n$. Une métrique riemannienne est alors donnée par la formule

$$ds^2 = \sum_{i,j=1}^{n} a_{ij}(q)\,dq_i dq_j, \quad a_{ij} = a_{ji},$$

où dq_i sont les coordonnées d'un vecteur tangent.

Les fonctions $a_{ij}(q)$ sont bien entendu supposées autant de fois différentiables qu'il le faut.

F. Dérivée d'une application. Soit $f\colon M\to N$ une application d'une variété M dans une variété N. L'application f est *différentiable* si elle est définie par des fonctions différentiables dans les coordonnées locales sur M et N.

D é f i n i t i o n. On appelle *dérivée* de l'application différentiable $f\colon M\to N$ au point $x\in M$ l'application linéaire des espaces tangents

$$f_{*x}\colon TM_x\to TN_{f(x)},$$

définie de la manière suivante (fig. 64).

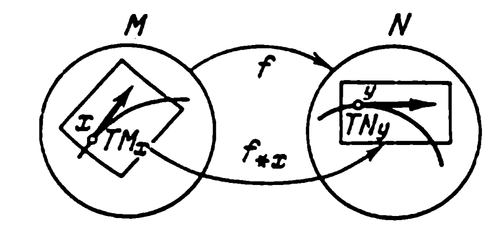


Fig. 64. Dérivée d'une application.

Soit $v\in TM_x$. Considérons une courbe $\varphi\colon \mathbf{R}\to M$, $\varphi(0)=x$ de vecteur vitesse $\left.\frac{d\varphi}{dt}\right|_{t=0}=v$. Alors $f_{*x}v$ est vecteur vitesse de la courbe $f\circ\varphi\colon \mathbf{R}\to N$,

$$f_{*x}v=\left.\frac{d}{dt}\right|_{t=0} f(\varphi(t)).$$

E x e r c i c e. Montrer que le vecteur $f_{*x}v$ dépend seulement du vecteur v et pas de la courbe φ.

E x e r c i c e. Montrer que l'application $f_{*x}: TM_x \to TN_{f(x)}$ est linéaire.

E x e r c i c e. Soient $x = (x_1, \ldots, x_m)$ des coordonnées au voisinage d'un point $x \in M$ et $y = (y_1, \ldots, y_n)$ des coordonnées au voisinage d'un point $y \in N$. Soient ξ une collection de composantes du vecteur v et η une collection de composantes du vecteur $f_{*x}v$. Démontrer que

$$\eta = \frac{\partial y}{\partial x}\xi, \text{ i.e. } \eta_i = \sum_j \frac{\partial y_i}{\partial x_j}\xi_j.$$

En groupant les applications f_{*x} pour tous les x, on obtient une application unique de tout le fibré tangent

$$f_*: TM \to TN, \quad f_*v = f_{*x}v \text{ pour } v \in TM_x.$$

E x e r c i c e. Montrer que f_* est une application différentiable.

E x e r c i c e. Soit $f: M \to N$, $g: N \to K$, $h = g \circ f: M \to K$. Montrer que $h_* = g_* \circ f_*$.

§ 19. Système dynamique lagrangien

Dans ce paragraphe on définit un système dynamique lagrangien sur une variété. Le système à liaisons holonomes en est un cas particulier.

A. Définition d'un système lagrangien. Soient M une variété différentiable, TM son fibré tangent, $L: TM \to \mathbf{R}$ une fonction différentiable. On dit de l'application $\gamma: \mathbf{R} \to M$ qu'elle est un *mouvement dans un système lagrangien de variété de configuration M et de lagrangien L* si γ est extrémale de la fonctionnelle

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(\dot{\gamma})\,dt,$$

où $\dot{\gamma}$ est le vecteur vitesse, $\dot{\gamma}(t) \in TM_{\gamma(t)}$.

E x e m p l e. Soit M un domaine d'un espace arithmétique muni des coordonnées $q = (q_1, \ldots, q_n)$. Le lagrangien $L: TM \to \mathbf{R}$ s'écrit sous forme d'une fonction de $2n$ coordonnées: $L(q, \dot{q})$. On a montré au § 12 que les variations des coordonnées du point en mouvement en fonction du temps satisfaisaient les équations de Lagrange.

Corollaire. *Les variations des coordonnées locales $q = (q_1, \ldots, q_n)$ d'un point $\gamma(t)$ en mouvement dans un système lagrangien sur une variété M vérifient les équations de Lagrange*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q},$$

où $L(q, \dot{q})$ est l'expression de la fonction $L: TM \to \mathbf{R}$ dans les coordonnées q et $\dot{q}$ sur TM.

Le cas particulier suivant se rencontre très fréquemment.

B. Système naturel. Soit M une variété riemannienne. On appelle *énergie cinétique* une fonction de vecteur tangent qui est une forme quadratique sur chaque espace tangent

$$T = \frac{1}{2}\langle v, v\rangle, \quad v \in TM_x.$$

On appelle *énergie potentielle* une fonction différentiable $U: M \to \mathbf{R}$.

D é f i n i t i o n. Un système lagrangien sur une variété riemannienne est *naturel* si le lagrangien est égal à la différence entre l'énergie cinétique et l'énergie potentielle: $L = T - U$.

E x e m p l e. Considérons deux points de masses m_1 et m_2 reliés par un segment de longueur l dans un plan x, y. La variété de configuration de dimension trois

$$M = \mathbf{R}^2 \times S^1 \subset \mathbf{R}^2 \times \mathbf{R}^2$$

est définie dans l'espace de configuration $\mathbf{R}^2 \times \mathbf{R}^2$ des deux points libres (x_1, y_1) et (x_2, y_2) par la condition $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = l$ (fig. 65).

La forme quadratique

$$m_1(\dot{x}_1^2 + \dot{y}_1^2) + m_2(\dot{x}_2^2 + \dot{y}_2^2)$$

est définie sur un espace tangent à l'espace $\{(x_1, x_2, y_1, y_2)\}$. Notre variété de dimension trois en tant que plongée dans une variété de dimension quatre est



Fig. 65. Segment d'un plan.

munie d'une métrique riemannienne. Le système holonome obtenu s'appelle en mécanique segment de longueur constante sur le plan x, y. Son énergie cinétique est donnée par la formule

$$T = m_1 \frac{\dot{x}_1^2 + \dot{y}_1^2}{2} + m_2 \frac{\dot{x}_2^2 + \dot{y}_2^2}{2}.$$

C. Système comportant des liaisons holonomes. Au § 17 nous avons défini un système de points matériels soumis à des liaisons holonomes. Montrons que ce système est naturel.

En effet, supposons la variété de configuration M d'un système à liaisons plongée dans l'espace de configuration d'un système de n points libres. Définissons une métrique de cet espace par la forme quadratique

$$\sum_{i=1}^{n} m_i \dot{r}_i^2.$$

Alors le système naturel correspondant à la variété riemannienne plongée M et à l'énergie potentielle U coïncide avec le système défini au § 17, ou avec le cas limite d'un système dont le potentiel $U + Nq_2^2$, $N \to \infty$, croît vite en dehors de M.

D. Marche à suivre pour résoudre les problèmes comportant des liaisons.

1. Trouver la variété de configuration et la munir des coordonnées $q_1, \ldots, q_k$ (d'une façon générale, locales dans le voisinage de chaque point).

2. Exprimer l'énergie cinétique $T = \sum \frac{1}{2} m_i \dot{r}_i^2$ par une forme quadratique en les vitesses généralisées

$$T = \frac{1}{2} \sum a_{ij}(q)\, \dot{q}_i \dot{q}_j.$$

3. Composer le lagrangien $L = T - U(q)$ et résoudre l'équation de Lagrange.

Exemple. Considérons le mouvement d'un point matériel de masse unité sur une surface de révolution dans l'espace ordinaire de dimension 3. On démontre que les orbites sont les géodésiques de la surface. En coordonnées cylindriques r, φ, z la surface est donnée (localement) sous la forme $r = r(z)$ ou $z = z(r)$. L'énergie cinétique s'écrit donc (fig. 66)

$$T = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}[(1 + r_z'^2)\,\dot{z}^2 + r^2(z)\,\dot{\varphi}^2] \text{ dans les coordonnées } \varphi, z;$$

$$= \frac{1}{2}[(1 + z_r'^2)\,\dot{r}^2 + r^2\dot{\varphi}^2] \text{ dans les coordonnées } r, \varphi.$$

(On s'est servi de la relation $\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2\dot{\varphi}^2$).

Le lagrangien $L = T$. Dans les deux systèmes la coordonnée φ est cyclique. L'impulsion correspondante se conserve; $p_\varphi = r^2\dot{\varphi}$ n'est autre que la z-ème

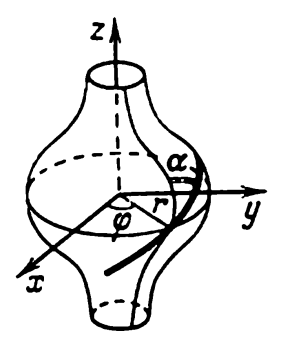



Fig. 66. Surface de révolution.

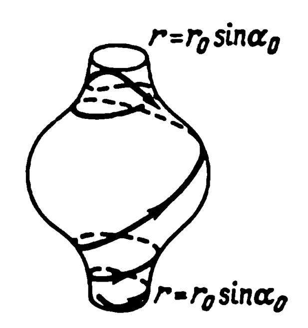


Fig. 67. Géodésiques sur une surface de révolution.

composante du moment de la quantité de mouvement. Comme le système possède deux degrés de liberté, la connaissance de la coordonnée φ suffit à intégrer complètement le problème (voir le corollaire 3 du § 15).

Pour avoir une idée exacte de l'allure des orbites il est plus simple de raisonner autrement. Désignons par α l'angle d'une orbite avec un méridien. On a $r\dot{\varphi} = |v| \sin \alpha$, où $|v|$ est le module du vecteur vitesse (fig. 66).

Or d'après la loi de conservation de l'énergie $H = L = T$ se conserve. Donc $|v| = \text{const}$. Et la loi de conservation de p_φ s'écrit

$$r \sin \alpha = \text{const}$$

(théorème de Clairault).

Cette relation montre que le mouvement s'effectue dans le domaine $|\sin \alpha| < 1$, i.e. $r \geqslant r_0 \sin \alpha_0$. En outre la pente de l'orbite par rapport au méridien s'accroît lorsque le rayon r décroît. Dès qu'elle atteint le plus petit $r = r_0 \sin \alpha_0$, l'orbite est réfléchie et revient dans le domaine des plus grands r (fig. 67).

Exercice. Montrer que toutes les géodésiques représentées sur la surface de révolution se répartissent en trois classes: méridiens, courbes fermées et géodésiques partout denses dans la couronne $r \geqslant c$.

Exercice. Etudier le comportement des géodésiques sur la surface du tore $(r - R)^2 + z^2 = \rho^2$.

E. Systèmes non autonomes. Un *système non autonome lagrangien* se distingue du système autonome que nous avons étudié jusqu'ici par le fait que le lagrangien dépend en plus du temps:

$$L: TM \times \mathbf{R} \to \mathbf{R}, \quad L = L(q, \dot{q}, t).$$

En particulier, dans un système non autonome naturel l'énergie cinétique comme l'énergie potentielle sont susceptibles de dépendre du temps:

$$T: TM \times \mathbf{R} \to \mathbf{R},\ U: M \times \mathbf{R} \to \mathbf{R},\ T = T(q, \dot{q}, t),\ U = U(q, t).$$

Un système de n points matériels comportant des liaisons holonomes fonctions du temps se définit à l'aide d'une sous-variété, variable avec le temps, de l'espace de configuration d'un système libre. Une telle variété est donnée par l'application

$$i: M \times \mathbf{R} \to E^{3n}, \quad i(q, t) = x,$$

qui pour tout $t \in \mathbf{R}$ fixe définit le plongement $M \to E^{3n}$. La marche à suivre du point D est valable pour les systèmes non autonomes.

Exemple. Le mouvement d'une perle sur un cercle vertical de rayon r (fig. 68) tournant à une vitesse angulaire ω autour d'un axe vertical passant par son centre O. La variété M est le cercle. Désignons par q la coordonnée angulaire sur le cercle: q sera mesurée à partir du point supérieur.

Soient x, y, z des coordonnées cartésiennes dans E^3, d'origine O et d'axe vertical z. Soit φ l'angle du plan du cercle avec le plan xOz. Par hypothèse $\varphi = \omega t$. L'application $i: M \times \mathbf{R} \to E^3$ est définie par la formule

$$i(q, t) = (r \sin q \cos \omega t,\ r \sin q \sin \omega t,\ r \cos q).$$

De cette formule (ou tout simplement d'un triangle rectangle infiniment petit) l'on déduit que

$$T = \frac{m}{2}(\omega^2 r^2 \sin^2 q + r^2 \dot{q}^2), \quad U = mgr \cos q.$$

Par bonheur le lagrangien $L = T - U$ ne dépend pas de t bien que les liaisons en dépendent. De plus il est le même que dans un système de dimension 1 d'énergie cinétique

$$T_0 = \frac{M}{2}\dot{q}^2, \quad M = mr^2,$$

et d'énergie potentielle

$$V = A \cos q - B \sin^2 q, \quad A = mgr, \quad B = \frac{m}{2}\omega^2 r^2.$$

L'allure des courbes de phase dépend des relations entre A et B. Si $2B < A$, i.e. le cercle tourne si lentement que $\omega^2 r < g$, la position inférieure de la perle

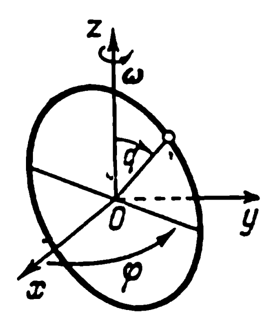


Fig. 68. Perle sur un cercle en rotation.

($q = \pi$) est stable et le caractère du mouvement est en général le même que dans le cas du pendule mathématique ($\omega = 0$).

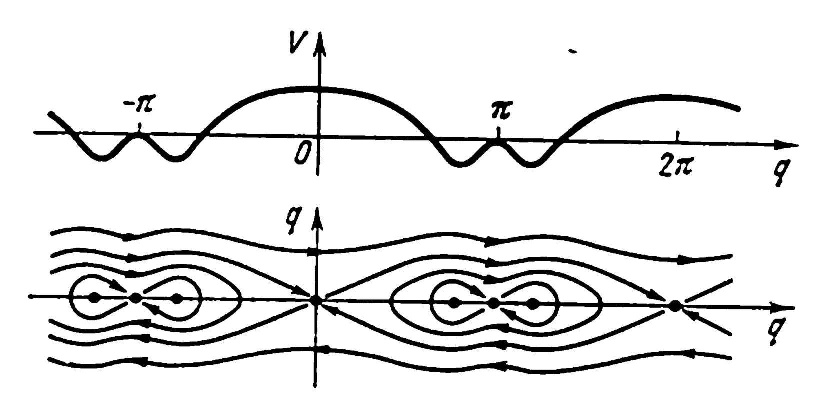


Fig. 69. Energie potentielle effective et plan de phase de la perle.

Si $2B > A$, i.e. la rotation du cercle est assez rapide, la position inférieure de la perle devient instable, par contre on aura deux positions stables sur le cercle: $\cos q = -\frac{A}{2B} = -\frac{g}{\omega^2 r}$. L'allure des courbes de phase représentées sur le plan $\{(q, \dot{q})\}$ (fig. 69) nous montre clairement comment va se comporter la perle pour différentes conditions initiales.

§ 20. Théorème de E. Noether

Les diverses lois de conservation (de l'impulsion, du moment, etc.) sont des cas particuliers d'un même théorème général : à tout groupe à un paramètre de difféomorphismes de la variété de configuration d'un système lagrangien, conservant la fonction de Lagrange, correspond une intégrale première des équations du mouvement.

A. Enoncé du théorème. Soit M une variété différentiable; L: $TM \to \mathbf{R}$ une fonction numérique différentiable sur son fibré tangent TM; h: $M \to M$ une application différentiable.

D é f i n i t i o n. Le système lagrangien (M, L) *admet l'application* h si pour tout vecteur tangent $v \in TM$

$$L(h_* v) = L(v).$$

E x e m p l e. Supposons que $M = \{(x_1, x_2, x_3)\}$, $L = \frac{m}{2}(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) - U(x_2, x_3)$. Le système admet la translation $h: (x_1, x_2, x_3) \to (x_1 + s, x_2, x_3)$ le long de l'axe x_1 et n'admet pas, d'une façon générale, de translation le long de l'axe x_2.

Théorème de Noether. *Si le système (M, L) admet un groupe à un paramètre de difféomorphismes h^s: $M \to M$, $s \in \mathbf{R}$, $h^0 = E$, alors le système d'équations de Lagrange correspondant à L a pour intégrale première $I: TM \to \mathbf{R}$.*

Dans des coordonnées locales q sur M l'intégrale I s'écrit

$$I(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \frac{dh^s(q)}{ds}\bigg|_{s=0}.$$

B. Démonstration. Supposons tout d'abord que $M = \mathbf{R}^n$ est un espace arithmétique. Soit $\varphi: \mathbf{R} \to M$, $q = \varphi(t)$ une solution des

Fig. 70. Sur le théorème de Noether.

équations de Lagrange. Comme h_* conserve L, alors la *solution translatée* $h^s \circ \varphi$: $\mathbf{R} \to M$ *quel que soit s est également solution des équations de Lagrange* *). Considérons l'application Φ: $\mathbf{R} \times \mathbf{R} \to \mathbf{R}^n$, $q = \Phi(s, t) = h^s(\varphi(t))$ (fig. 70).

*) Certains auteurs affirment à tort que la réciproque est vraie, i. e. si h^s envoie solutions dans solutions, alors h_*^s conserve L.

Affectons les dérivées par rapport à t d'un point et les dérivées par rapport à s du symbole prime. Par hypothèse

$$0=\frac{\partial L(\Phi, \dot{\Phi})}{\partial s}=\frac{\partial L}{\partial \dot{q}}\Phi'+\frac{\partial L}{\partial q}\dot{\Phi}', \tag{1}$$

où les dérivées partielles de L sont prises au point $q = \Phi(s, t)$, $\dot{q} = \dot{\Phi}(s, t)$.

Nous avons souligné plus haut que quel que soit s fixe l'application $\Phi|_{s=\text{const}}: \mathbf{R} \to \mathbf{R}^n$ satisfait à l'équation de Lagrange

$$\frac{\partial}{\partial t}\left[\frac{\partial L}{\partial \dot{q}}(\Phi(s, t), \dot{\Phi}(s, t))\right]=\frac{\partial L}{\partial q}(\Phi(s, t), \dot{\Phi}(s, t)).$$

Introduisons la notation $F(s, t) = \frac{\partial L}{\partial \dot{q}}(\Phi(s, t), \dot{\Phi}(s, t))$ et remplaçons $\frac{\partial L}{\partial q}$ par $\frac{\partial F}{\partial t}$ dans (1).

En écrivant $\dot{q}'$ sous la forme $\frac{d}{dt}q'$ on obtient

$$0=\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)q'+\frac{\partial L}{\partial \dot{q}}\left(\frac{d}{dt}q'\right)=\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}q'\right)=\frac{dI}{dt}, \text{ c.q.f.d.}$$

R e m a r q u e. L'intégrale première $I = \frac{\partial L}{\partial \dot{q}}q'$ a été définie plus haut à l'aide des coordonnées locales q. Il s'avère que la quantité $I(v)$ *ne dépend pas du choix du système de coordonnées* q.

En effet, I est la vitesse de variation de $L(v)$ lorsque le vecteur $v \in TM_x$ varie à l'intérieur de TM_x à la vitesse $\left.\frac{d}{ds}\right|_{s=0}h^s_*x$. Donc $I(v)$ est une fonction intrinsèquement définie, dans son ensemble, du vecteur tangent $v \in TM_x$. Ce qui démontre le théorème de Noether dans le cas où M est une variété.

C. Exemples.

E x e m p l e 1. Considérons un système de points matériels de masses m_i:

$$L=\sum m_i\frac{\dot{x}_i^2}{2}-U(x), \quad x_i=x_{i1}e_1+x_{i2}e_2+x_{i3}e_3,$$

comportant des liaisons $f_j(x) = 0$. Supposons que ce système admet des translations le long de l'axe e_1:

$$h^s: x_i \to x_i + se_1 \ \forall i.$$

En d'autres termes, les liaisons permettent au système de se déplacer globalement le long de l'axe e_1 sans que l'énergie potentielle change.

Le théorème de Noether entraîne:

Si le système admet des translations le long de l'axe e_1, alors la projection de son centre d'inertie sur l'axe e_1 est en mouvement rectiligne uniforme.

En effet $\frac{d}{ds}\Big|_{s=0} h^s x_i = e_1$. D'après la remarque faite à la fin du point B, se conserve la grandeur

$$I = \sum \frac{\partial L}{\partial \dot{x}_i} e_1 = \sum m_i \dot{x}_{i1},$$

i.e. la première composante P_1 du vecteur impulsion. Pour les systèmes sans liaisons nous avons démontré cela avant.

E x e m p l e 2. *Si le système admet des rotations autour de l'axe e_1, alors se conserve le moment cinétique par rapport à cet axe*

$$M_1 = \sum_i ([x_i, m_i \dot{x}_i], e_1).$$

En effet, on s'assure sans peine que si h^s est une rotation d'angle s autour de l'axe e_1, alors

$$\frac{d}{ds}\Big|_{s=0} h^s x_i = [e_1, x_i], \text{ d'où } I = \sum_i \frac{\partial L}{\partial \dot{x}_i} [e_1, x_i] =$$

$$= \sum_i (m_i \dot{x}_i, [e_1, x_i]) = \sum_i ([x_i, m_i \dot{x}_i], e_1).$$

E x e r c i c e 1. Soit une particule en mouvement dans le champ d'une ligne hélicoïdale homogène $x = \cos \varphi$, $y = \sin \varphi$, $z = c\varphi$. Trouver la loi de conservation correspondant à cette symétrie hélicoïdale.

Réponse. Dans tout système admettant des mouvements hélicoïdaux laissant en place notre ligne hélicoïdale se conserve la grandeur $I = cP_3 + M_3$.

E x e r c i c e 2. Soit un solide se déplaçant par inertie. Montrer que son centre d'inertie est en mouvement rectiligne uniforme. Si le centre d'inertie est au repos, le moment cinétique par rapport à lui se conserve.

E x e r c i c e 3. Quelles sont les grandeurs qui se conservent dans le mouvement d'un solide pesant ayant un point fixe O? Etudier le cas où le solide est symétrique par rapport à un axe passant par O.

E x e r c i c e 4. Généraliser le théorème de Noether à un système lagrangien non autonome.

I n d i c a t i o n. Soit $M_1 = M \times \mathbf{R}$ la variété de configuration élargie (le produit cartésien de la variété de configuration M par l'axe temporel $\mathbf{R}$).

Définissons la fonction $L_1 \colon TM_1 \to \mathbf{R}$ comme $L\, dt/d\tau$, i.e. définissons-la par la formule

$$L_1\left(q, t, \frac{dq}{d\tau}, \frac{dt}{d\tau}\right) = L\left(q, \frac{dq}{d\tau}\Big/\frac{dt}{d\tau}, t\right)\frac{dt}{d\tau}$$

dans des coordonnées locales q et t sur M_1. Appliquons le théorème de Noether au système lagrangien (M_1, L_1).

Si L_1 admet les transformations $h^s: M_1 \to M_1$, nous trouvons une intégrale première $I_1: TM_1 \to \mathbf{R}$. Comme $\int L\,dt = \int L_1\,d\tau$, nous sommes conduits à une intégrale première $I: TM \times \mathbf{R} \to \mathbf{R}$ du système de départ. Si dans les coordonnées locales (q, t) sur M_1 on a $I_1 = I_1\left(q, t, \frac{dq}{d\tau}, \frac{dt}{d\tau}\right)$, alors $I(q, \dot{q}, t) = I_1(q, t, \dot{q}, 1)$.

En particulier, si L ne dépend pas explicitement du temps, alors L_1 admet les translations dans le temps $h^s(q, t) = (q, t + s)$. L'intégrale première correspondante I est l'intégrale d'énergie.

§ 21. Principe de D'Alembert

On donne ici une nouvelle définition d'un système de points matériels comportant des liaisons holonomes et on démontre qu'elle est équivalente à celle du § 17.

A. Exemple. Considérons un système holonome (M, L) où M est une surface d'un espace $\{x\}$ de dimension 3:

$$L = \frac{1}{2} m\dot{x}^2 - U(x)$$

ou en termes de mécanique: « le point matériel x de masse m est astreint à se déplacer sur la surface différentiable M ».

Considérons le mouvement $x(t)$ d'un point. Si l'équation de Newton $m\ddot{x} + \frac{\partial U}{\partial x} = 0$ était vérifiée, alors en l'absence de forces extérieures ($U = 0$), la trajectoire serait une droite et ne pourrait donc pas être située sur la surface M.

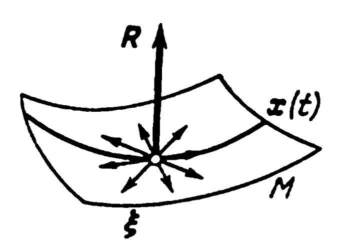


Fig. 71. Réaction d'une liaison.

Du point de vue newtonien cela implique l'existence d'une nouvelle force « astreignant le point à rester sur la surface ».

D é f i n i t i o n. La quantité

$$\boldsymbol{R} = m\ddot{x} + \frac{\partial U}{\partial x}$$

s'appelle *réaction de la liaison* (fig. 71).

Compte tenu de la réaction $\boldsymbol{R}(t)$, les équations de Newton s'écrivent visiblement

$$m\ddot{x} = -\frac{\partial U}{\partial x} + \boldsymbol{R}.$$

Le sens physique de la réaction R devient clair si l'on considère notre système lié comme la limite de systèmes d'énergie potentielle $U + NU_1$, $N \to \infty$, $U_1(x) = \rho^2(x, M)$. Pour N grand le potentiel de liaison NU_1 conditionne la force à variation rapide $F = -N\frac{\partial U_1}{\partial x}$. Lorsqu'on passe à la limite ($N \to \infty$), il reste la valeur moyenne R de F pendant les oscillations de x près de M. La force F est orthogonale à la surface M. Donc la réaction R est orthogonale à M: $(R, \xi) = 0$ pour tout vecteur tangent ξ.

B. Enoncé du principe de D'Alembert-Lagrange. En mécanique les vecteurs tangents à la variété de configuration s'appellent *déplacements virtuels*. Le principe de D'Alembert-Lagrange dit:

$$\left(m\ddot{x} + \frac{\partial U}{\partial x},\ \xi\right) = 0$$

pour tout déplacement virtuel ξ ou encore: *le travail de la force de réaction est nul pour n'importe quel déplacement virtuel.*

Dans le cas d'un système de points matériels x_i de masses m_i les forces de réaction R_i sont définies comme $R_i = m_i\ddot{x}_i + \frac{\partial U}{\partial x_i}$ et le principe de D'Alembert est de la forme $\sum(R_i, \xi_i) = 0$ ou $\sum\left(\left(m_i\ddot{x}_i + \frac{\partial U}{\partial x_i}\right),\ \xi_i\right) = 0$, i.e. *la somme des travaux des forces de réaction est nulle pour n'importe quel déplacement virtuel* $\{\xi_i\} \in TM_x$.

Les liaisons possédant la propriété indiquée ci-dessus sont dites *parfaites.*

Si l'on définit un système soumis à une liaison holonome comme une limite comprise au sens ci-dessus lorsque $N \to \infty$, alors le principe de D'Alembert-Lagrange a force de théorème; sa démonstration a été ébauchée plus haut pour un cas simple.

On peut toutefois *définir* une liaison holonome parfaite à l'aide du principe de D'Alembert-Lagrange.

Nous avons donc trois définitions d'un système holonome lié:

1) La limite de systèmes d'énergie potentielle $U + NU_1$, $N \to \infty$.

2) Un système holonome (M, L) où M est une sous-variété de l'espace de configuration d'un système sans liaisons et L le lagrangien.

3) Un système vérifiant le principe de D'Alembert-Lagrange.

Ces trois définitions sont mathématiquement équivalentes.

La démonstration des assertions 1) $\Longrightarrow$ 2) et 1) $\Longrightarrow$ 3) a été tracée plus haut et ne sera pas faite dans le détail. Nous montrerons que 2) $\Longleftrightarrow$ 3).

C. Equivalence du principe de D'Alembert-Lagrange et du principe variationnel. Soit M une sous-variété d'un espace euclidien $M \subset \mathbf{R}^N$ et x: $\mathbf{R} \to M$ une courbe; $x(t_0) = x_0$, $x(t_1) = x_1$.

Définition. On dit qu'une courbe x est *extrémale liée* de la fonctionnelle d'action

$$\Phi = \int_{t_0}^{t_1} \left\{\frac{\dot{x}^2}{2} - U(x)\right\} dt,$$

si la différentielle $\delta\Phi = 0$ *sous réserve* que pour la comparaison l'on prenne des courbes voisines *) passant par x_0 et x_1 sur M.

On écrira

$$\delta_M \Phi = 0. \tag{1}$$

Il est évident que l'équation (1) est équivalente aux équations de Lagrange

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}, \quad L = \frac{\dot{x}^2}{2} - U(x), \quad x = x(q),$$

dans chaque système de coordonnées locales q sur M.

Théorème. *Une condition nécessaire et suffisante pour qu'une courbe* $x\colon \mathbf{R} \to M \subset \mathbf{R}^N$ *soit extrémale liée de l'action (i.e. vérifie l'équation* (1)) *est qu'elle vérifie l'équation de D'Alembert*

$$\left(\ddot{x} + \frac{\partial U}{\partial x}, \xi\right) = 0, \quad \forall \xi \in TM_x. \tag{2}$$

Lemme. *Soit* $f\colon \{t\colon t_0 \leqslant t \leqslant t_1\} \to \mathbf{R}^N$ *un champ de vecteurs continu. Si pour tout champ de vecteurs* ξ *continu et tangent à* M *le long de* x ($\xi(t) \in TM_{x(t)}$, $\xi(t)$ *s'annule pour* $t = t_0, t_1$) *l'on a*

$$\int_{t_0}^{t_1} f(t)\,\xi(t)\,dt = 0,$$

alors le champ $f(t)$ *est en chaque point de* $x(t)$ *orthogonal à la surface* M *(i.e.* $(f(t), h) = 0$ *pour tout vecteur* $h \in TM_{x(t)}$) (fig. 72).

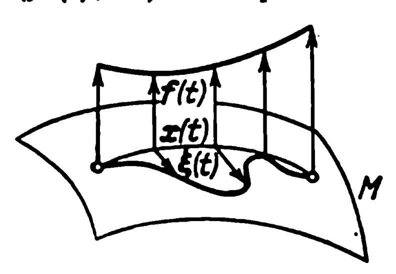


Fig. 72. Lemme du champ normal.

La démonstration du lemme reprend les raisonnements qui nous ont servi à déduire les équations d'Euler-Lagrange au § 12.

D é m o n s t r a t i o n d u t h é o r è m e. Comparons les valeurs de Φ sur les courbes $x(t)$ et $x(t) + \xi(t)$, $\xi(t_0) = \xi(t_1) = 0$. Une intégration par parties donne

$$\delta\Phi = \int_{t_0}^{t_1} \left(\dot{x}\dot{\xi} - \frac{\partial U}{\partial x}\xi\right) dt = -\int_{t_0}^{t_1} \left(\ddot{x} + \frac{\partial U}{\partial x}\right) \xi\, dt.$$

*) En toute rigueur, pour définir la variation $\delta\Phi$ il faudrait définir la structure du domaine de l'espace vectoriel dans l'ensemble des courbes proches de x sur M. On peut le faire à l'aide des coordonnées sur M; en outre la propriété d'être extrémale liée ne dépend pas du choix du système de coordonnées.

Cette formule entraîne *) l'équivalence de l'équation (1) $\delta_M \Phi = 0$ et de l'ensemble des équations

$$\int_{t_0}^{t_1} \left(\ddot{x} + \frac{\partial U}{\partial x} \right) \xi \, dt = 0 \tag{3}$$

pour tous les champs vectoriels tangents $\xi(t) \in TM_{x(t)}$, $\xi(t_0) = \xi(t_1) = 0$. En vertu du lemme $\left(\text{dans lequel on fera } f = \ddot{x} + \frac{\partial U}{\partial x}\right)$ l'ensemble des équations (3) est équivalent à l'équation de D'Alembert-Lagrange (2), c.q.f.d.

D. Remarques.

Remarque 1. Déduisons du théorème démontré le *principe de D'Alembert-Lagrange pour un système de n points* $x_i \in \mathbf{R}^3$, $i = 1, \ldots, n$ de masses m_i comportant des liaisons holonomes.

Dans les coordonnées $\bar{x} = \{\bar{x}_i = \sqrt{m_i}\, x_i\}$ l'énergie cinétique prend la forme $T = \frac{1}{2} \sum m_i \dot{x}_i^2 = \frac{1}{2} \dot{\bar{x}}^2$.

D'après le théorème démontré plus haut, les extrémales du principe de moindre action réalisent la condition

$$\left(\ddot{\bar{x}} + \frac{\partial U}{\partial \bar{x}}, \bar{\xi} \right) = 0$$

(principe de D'Alembert-Lagrange) pour un point de $\mathbf{R}^{3n}$: la force de réaction de dimension $3n$ est orthogonale à la variété M dans la métrique T). Revenons aux coordonnées x_i:

$$0 = \left(\sqrt{m_i}\, \ddot{x}_i + \frac{\partial U}{\partial \sqrt{m_i}\, x_i},\ \sqrt{m_i}\, \xi_i \right) = \sum_i \left(m_i \ddot{x}_i + \frac{\partial U}{\partial x_i},\ \xi_i \right),$$

i.e. on obtient le principe de D'Alembert-Lagrange sous sa forme précédente: la somme des travaux des forces de réaction est nulle pour un déplacement virtuel.

Remarque 2. Le principe de D'Alembert-Lagrange peut s'énoncer sous une forme un peu différente si l'on fait intervenir la statique. On appelle *position d'équilibre* un point x_0 tel qu'il soit orbite du mouvement: $x(t) \equiv x_0$.

Supposons qu'un point matériel se déplace sur une surface différentiable M sous l'action d'une force $f = -\partial U / \partial x$.

Théorème. *Pour qu'un point x_0 de la surface M soit position d'équilibre il est nécessaire et suffisant que la force f soit orthogonale à M: $(f(x_0), \xi) = 0$ pour tous les $\xi \in TM_{x_0}$.*

---

*) La distance du point $x(t) + \xi(t)$ à M est du deuxième ordre de petitesse par rapport à $\xi(t)$.

7*

Ceci découle de l'équation de D'Alembert-Lagrange puisque $x = 0$.

D é f i n i t i o n. La quantité $-m\ddot{x}$ est appelée *force d'inertie*. Le principe de D'Alembert-Lagrange s'énonce maintenant:

Théorème. *Si à l'action de toutes les forces on ajoute celle de la force d'inertie, alors x devient position d'équilibre.*

En effet, l'équation de D'Alembert

$$(-m\ddot{x} + f,\ \xi) = 0$$

traduit, en vertu du théorème précédent, le fait que x est position d'équilibre du système sollicité par les forces $-m\ddot{x} + f$.

Ce qu'on vient de démontrer se généralise à un système de points:

Si $x = \{x_i\}$ est position d'équilibre, la somme des travaux des forces agissant sur ce système est nulle pour un déplacement virtuel.

Si aux forces agissant sur ce système on ajoute la force d'inertie $-m_i\ddot{x}_i(t)$, la position $x(t)$ deviendra position d'équilibre.

Donc le problème du mouvement nous ramène à un problème d'équilibre sous l'action d'autres forces.

R e m a r q u e 3. Jusqu'ici nous n'avons pas étudié les cas où les liaisons dépendaient du temps. Tout ce qui a été dit plus haut s'étend sans changement à de telles liaisons.

E x e m p l e. Soit une perle glissant sur une tige faisant un angle α avec un axe vertical et animée d'une rotation uniforme de vitesse

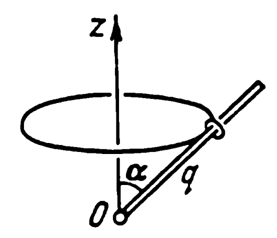


Fig. 73. Perle sur une tige en rotation.

angulaire ω autour de cet axe (on néglige le poids de la perle). Prenons pour coordonnée q la distance au point O (fig. 73). L'énergie cinétique et le lagrangien ont pour expression:

$$L = T = \frac{1}{2} mv^2 = \frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 r^2, \quad r = q \sin \alpha.$$

L'équation de Lagrange s'écrit: $m\ddot{q} = m\omega^2 q \sin^2 \alpha$.

La force de réaction est à chaque instant orthogonale aux déplacements virtuels (i.e. à la direction de la tige), mais pas à la trajectoire réelle.

R e m a r q u e 4. De l'équation de D'Alembert-Lagrange on déduit sans peine les lois de conservation. Si par exemple les déplacements virtuels renferment une translation le long de l'axe x_1

$$\xi_i = e_1,$$

alors la somme des travaux des forces de réaction est nulle pour ce déplacement:

$$\sum (R_i, e_1) = (\sum R_i, e_1) = 0.$$

Maintenant nous allons considérer les forces de réaction comme des forces extérieures. On remarquera alors que la somme des premières composantes des forces extérieures est nulle. Cela signifie que la première composante P_1 du vecteur de la quantité de mouvement est invariante.

Nous avons déjà obtenu ce résultat à partir du théorème de Noether.

R e m a r q u e 5. Rappelons que l'holonomie de telle ou telle liaison physique (avec une plus ou moins grande précision) est une affaire d'expérience. Du point de vue mathématique, l'holonomie des liaisons est un postulat d'origine physique ; on peut l'adopter sous des formes équivalentes différentes, par exemple sous la forme du principe de moindre action (1) ou du principe de D'Alembert-Lagrange (2), mais quand on définit les liaisons, il s'agit toujours de faits d'expérience non contenus dans les équations de Newton.

R e m a r q u e 6. Notre terminologie est légèrement différente de celle adoptée dans les cours de mécanique où le principe de D'Alembert-Lagrange s'étend à une classe plus vaste de systèmes (systèmes non holonomes comportant des liaisons parfaites). Nous n'étudierons pas les systèmes non holonomes dans le cadre de cet ouvrage. Nous remarquerons simplement qu'un exemple de système non holonome nous est fourni par une boule roulant sans glissement sur un plan. En tout point d'un espace tangent à la variété de configuration d'un système non holonome est fixé un sous-espace auquel doit être tangent le vecteur vitesse.

R e m a r q u e 7. Si le système est constitué de points matériels reliés par des tiges, des charnières, etc., on peut être tenté de parler de la réaction de telle ou telle liaison.

Nous avons défini la résultante R_i des forces de réaction de toutes les liaisons pour chaque point matériel m_i. La notion de force de réaction d'une liaison isolée est *impossible* à définir comme il ressort de l'exemple simple d'une poutre soutenue par trois colonnes. Si l'on essaye de définir les forces de réaction R_1, R_2, R_3 des colonnes par passage à la limite (en assimilant les colonnes à des ressorts très rigides), on s'assure que le résultat dépend de la distribution de la rigidité.

Dans les recueils d'exercices les problèmes sont choisis de façon à s'affranchir de cette difficulté.

E x e r c i c e. Une tige de poids P, faisant un angle de 60° avec le plan d'une table, commence à tomber sans vitesse initiale (fig. 74). Trouver la force

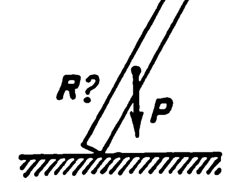


Fig. 74. Force de réaction de la table.

de réaction de la table à l'instant initial en supposant que la table est a) absolument lisse, b) absolument rugueuse. (Dans le premier cas, la liaison holonome maintient le bout de la tige sur le plan de la table, dans le deuxième, au point donné.)

CHAPITRE 5

OSCILLATIONS

Les équations linéaires étant faciles à résoudre, la théorie des oscillations linéaires est le chapitre le mieux élaboré de la mécanique. Dans nombre de problèmes non linéaires, la linéarisation fournit une solution approchée satisfaisante. Cependant même s'il n'en est pas ainsi, l'étude du problème linéarisé est souvent le premier pas que l'on entreprend pour vérifier la conformité des mouvements du système non linéaire avec son modèle linéaire.

§ 22. Linéarisation

On donne la définition des petites oscillations.

A. Positions d'équilibre.

D é f i n i t i o n. Un point x_0 est *position d'équilibre* du système

$$\frac{dx}{dt}=f(x), \quad x\in \mathbf{R}^n, \tag{1}$$

si $x(t)\equiv x_0$ est solution de ce système. En d'autres termes, $f(x_0)=0$, i.e. le champ de vecteurs $f(x)$ s'annule au point x_0.

E x e m p l e. Considérons un système dynamique naturel de lagrangien $L(q, \dot{q})=T-U$, $T=\frac{1}{2}\sum a_{ij}(q)\dot{q}_i\dot{q}_j\geqslant 0$, $U=U(q)$:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}=\frac{\partial L}{\partial q}, \quad q=(q_1, \ldots, q_n). \tag{2}$$

Les équations de Lagrange peuvent être écrites sous la forme d'un système de $2n$ équations du premier ordre (1). Essayons d'en trouver les positions d'équilibre.

Théorème. *Pour qu'un point $q=q_0$, $\dot{q}=\dot{q}_0$ soit position d'équilibre du système* (2) *il faut et il suffit que $\dot{q}_0=0$ et que ce point soit point critique de l'énergie potentielle, i.e.*

$$\left.\frac{\partial U}{\partial q}\right|_{q_0}=0. \tag{3}$$

D é m o n s t r a t i o n. Ecrivons les équations de Lagrange

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}} = \frac{\partial T}{\partial q} - \frac{\partial U}{\partial q}.$$

De (2) on voit que pour $\dot{q} = 0$ on aura $\frac{\partial T}{\partial q} = 0$, $\frac{\partial T}{\partial \dot{q}} = 0$. Donc $q = q_0$ est solution dans le cas (3) et dans ce cas seulement, c.q.f.d.

B. Stabilité des positions d'équilibre. Etudions maintenant les mouvements pour des conditions initiales proches d'une position d'équilibre.

Théorème. *Si un point q_0 est un minimum local strict de l'énergie potentielle U, alors la position d'équilibre $q = q_0$ est stable au sens de Liapounov.*

D é m o n s t r a t i o n. Soit $U(q_0) = h$. Pour $\varepsilon > 0$ suffisamment petit la composante connexe de l'ensemble $\{q : U(q) \leqslant h + \varepsilon\}$ qui contient q_0 sera un voisinage aussi petit que l'on veut du point q_0 (fig. 75). Ceci étant, la composante connexe du domaine correspondant $\{p, q : E(p, q) \leqslant h + \varepsilon\}$ $\left(p = \frac{\partial T}{\partial \dot{q}}\right.$ est l'impulsion, $E = T + U$ l'énergie totale$\left.\right)$ sera un voisinage aussi petit que l'on veut du point $p = 0$, $q = q_0$ dans l'espace des phases $\{(p, q)\}$.

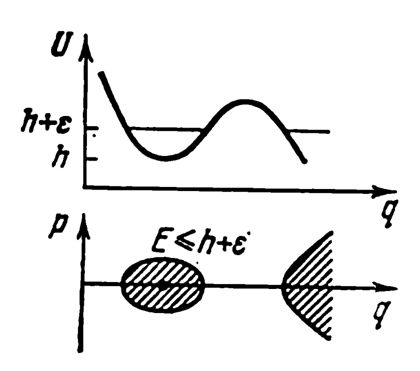


Fig. 75. Position d'équilibre stable.

Or le domaine $\{p, q : E \leqslant h + \varepsilon\}$ est invariant par le flot en vertu de la loi de conservation de l'énergie. Donc pour des conditions initiales $p(0)$, $q(0)$ suffisamment proches de $(0, q_0)$ la trajectoire de phase $p(t)$, $q(t)$ est tout entière proche de $(0, q_0)$, c.q.f.d.

E x e r c i c e. Est-ce que la position d'équilibre $q = q_0, p = 0$ peut être asymptotiquement stable?

E x e r c i c e. Démontrer que dans un système *analytique* à un degré de liberté, une position d'équilibre q_0 qui n'est pas un point de minimum local strict de l'énergie potentielle est instable au sens de Liapounov. Citer un exemple de système indéfiniment différentiable où ce n'est pas le cas.

R e m a r q u e. Tout laisse à croire que dans un système analytique à n degrés de liberté, une position d'équilibre qui n'est pas point de minimum est instable, mais ceci n'a pas été démontré.

C. Linéarisation d'une équation différentielle. Revenons maintenant au système général (1). On se sert souvent de la *linéarisation* pour étudier les solutions du système (1) proches d'une

position d'équilibre x_0. Supposons que $x_0 = 0$ (le cas général s'y ramène par un changement de coordonnées). Le premier terme de la série de Taylor de f est linéaire

$$f(x) = Ax + R_2(x), \quad A = \left.\frac{\partial f}{\partial x}\right|_0, \qquad R_2 = O(x^2),$$

où l'opérateur linéaire A dans les coordonnées $x_1, \ldots, x_n$ est associé à la matrice a_{ij}:

$$(Ax)_i = \sum_{ij} a_{ij} x_j; \qquad a_{ij} = \frac{\partial f_i}{\partial x_j}.$$

D é f i n i t i o n. On appelle *linéarisation* du système (1) le passage de ce système au système

$$\frac{dy}{dt} = Ay \quad (x \in \mathbf{R}^n, \quad y \in T\mathbf{R}^n_0). \tag{4}$$

E x e r c i c e. Montrer que la linéarisation est une opération intrinsèquement définie: l'opérateur A ne dépend pas du système de coordonnées.

Le système linéarisé a l'avantage d'être linéaire et par conséquent immédiatement résoluble

$$y(t) = e^{At} y(0), \quad \text{où } e^{At} = E + At + \frac{A^2 t^2}{2!} + \ldots$$

Sachant les solutions du système linéarisé (4) on peut se faire une idée des solutions du système de départ (1). Pour des x suffisamment petits l'écart $R_2(x)$ entre le système linéarisé et le système initial est petit devant x. C'est pourquoi les solutions $y(t)$ et $x(t)$ des deux systèmes vérifiant les conditions initiales $y(0) = x(0) = x_0$ restent voisines assez longtemps. Plus exactement on démontre facilement le théorème suivant:

Théorème. *Quels que soient $T > 0$ et $\varepsilon > 0$ il existe un $\delta > 0$ tel que si $|x(0)| < \delta$, alors $|x(t) - y(t)| < \varepsilon\delta$ pour tous les t de l'intervalle $0 < t < T$.*

D. Linéarisation d'un système lagrangien. Reprenons le système lagrangien (2) et essayons de la linéariser au voisinage d'une position d'équilibre $q = q_0$. Pour simplifier les formules plaçons l'origine des coordonnées en q_0.

Théorème. *Pour linéariser le système lagrangien (2) au voisinage de la position d'équilibre $q = 0$, il suffit de remplacer l'énergie cinétique $T = \frac{1}{2} \sum a_{ij}(q) \dot{q}_i \dot{q}_j$ par sa valeur pour $q = 0$*

$$T_2 = \frac{1}{2} \sum a_{ij} \dot{q}_i \dot{q}_j, \quad a_{ij} = a_{ij}(0),$$

et l'énergie potentielle $U(q)$ par sa partie quadratique

$$U_2 = \frac{1}{2}\sum b_{ij} q_i q_j, \quad b_{ij} = \left.\frac{\partial^2 U}{\partial q_i \partial q_j}\right|_{q=0}.$$

D é m o n s t r a t i o n. Ramenons le système de Lagrange à la forme (1) en se servant des variables canoniques p, q:

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}, \quad H(p, q) = T + U.$$

Comme $p = q = 0$ est position d'équilibre, les développements des seconds membres en série de Taylor au point 0 commencent par les termes linéaires en p et q. Comme les seconds membres sont des dérivées partielles, ces termes *linéaires* proviennent des termes quadratiques H_2 du développement de $H(p, q)$. Or H_2 est la fonction de Hamilton d'un système dont le lagrangien est $L_2 = T_2 - U_2$, puisque de toute évidence $H_2 = T_2(p) + U_2(q)$. Donc les équations linéarisées sont les équations de mouvement du système, de lagrangien $L_2 = T_2 - U_2$, décrit dans le théorème, c.q.f.d.

E x e m p l e. Considérons le système à un degré de liberté

$$T = \frac{1}{2} a(q)\dot{q}^2, \quad U = U(q).$$

Soit $q = q_0$ une position d'équilibre stable:

$$\left.\frac{\partial U}{\partial q}\right|_{q=q_0} = 0, \quad \left.\frac{\partial^2 U}{\partial q^2}\right|_{q=q_0} > 0 \text{ (fig. 76).}$$

Nous savons, de l'allure des courbes de phase, que pour des conditions initiales proches de $q = q_0$, $p = 0$ la solution est périodique

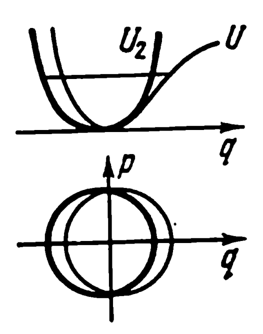


Fig. 76. Linéarisation.

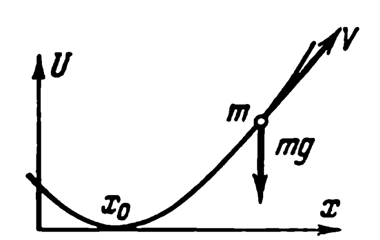


Fig. 77. Perle sur un fil.

de période τ dépendant en général des conditions initiales. Les deux théorèmes précédents entraînent le

Corollaire. *Au voisinage de la position d'équilibre q_0 la période τ d'oscillations tend vers la limite $\tau_0 = \frac{2\pi}{\omega_0}$, où $\omega_0^2 = \frac{b}{a}$,*

$$b = \frac{1}{2}\left.\frac{\partial^2 U}{\partial q^2}\right|_{q=q_0}, \quad a = a(q_0),$$

lorsque l'amplitude décroît.

En effet, $T_2 = \frac{1}{2} a\dot{q}^2$, $U_2 = \frac{1}{2} bq^2$ (on suppose $q_0 = 0$) pour le système linéarisé. Les solutions de l'équation de Lagrange $\ddot{q} = -\omega_0^2 q$ sont de période $\tau_0 = \frac{2\pi}{\omega_0}$,

$$q = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t$$

quelle que soit l'amplitude initiale.

E. Petites oscillations.

Définition. On appelle *petites oscillations* *) les mouvements d'un système linéarisé ($L_2 = T_2 - U_2$) au voisinage d'une position d'équilibre $q = q_0$. Dans le problème de dimension 1 les nombres τ_0, ω_0 sont appelés *période* et *amplitude des petites oscillations.*

Exemple. Trouver la période des petites oscillations d'une perle de masse unité enfilée sur un fil $y = U(x)$ dans un champ de pesanteur avec $g = 1$ au voisinage d'une position d'équilibre $x = x_0$ (fig. 77).

Solution. On a

$$U = mgy = U(x),$$

$$T = \frac{1}{2} mv^2 = \frac{1}{2}\left[1 + \left(\frac{\partial U}{\partial x}\right)^2\right].$$

Supposons que x_0 est une position d'équilibre stable $\left.\frac{\partial U}{\partial x}\right|_{x_0} = 0$; $\left.\frac{\partial^2 U}{\partial x^2}\right|_{x_0} > 0$. Alors la fréquence ω des petites oscillations est définie par la formule

$$\omega^2 = \left.\frac{\partial^2 U}{\partial x^2}\right|_{x_0},$$

car pour le système linéarisé $T_2 = \frac{1}{2}\dot{q}^2$, $U_2 = \frac{1}{2}\omega^2 q^2$ $(q = x - x_0)$.

Exercice. Montrer que non seulement les petites oscillations, mais le mouvement tout entier de la perle équivaut exactement au mouvement d'un certain système de dimension 1 de lagrangien

$$L = \frac{1}{2}\dot{q}^2 - V(q).$$

Indication. Prendre pour q la longueur sur le fil.

§ 23. Petites oscillations

On montre ici qu'un système lagrangien effectuant des petites oscillations se décompose en un produit direct de systèmes à un degré de liberté.

A. Problème du couple de formes. Penchons-nous plus en détail sur le problème des petites oscillations. Plus exactement, soit un système dont l'énergie cinétique et l'énergie potentielle sont les

*) Si la position d'équilibre est instable, nous parlerons de « petites oscillations instables », bien que le mouvement ne soit pas oscillatoire.

formes quadratiques

$$T=\frac{1}{2}(A\dot{q},\ \dot{q}),\quad U=\frac{1}{2}(Bq,\ q),\ q\in \mathbf{R}^n,\ \dot{q}\in \mathbf{R}^n. \tag{1}$$

L'énergie cinétique est une forme défínie positive.

Pour intégrer les équations de Lagrange il nous faut choisir convenablement les coordonnées.

Du cours d'algèbre linéaire on sait qu'*un couple de formes quadratiques* $(Aq,\ q)$, $(Bq,\ q)$, *dont la première est définie positive, peut être réduit à leurs axes principaux par un seul changement linéaire de coordonnées* *) :

$$Q = Cq,\quad Q = (Q_1,\ \ldots,\ Q_n).$$

Ceci étant, les coordonnées Q peuvent être choisies de sorte que la forme $(Aq,\ q)$ se réduise à une somme de carrés $(Q,\ Q)$. Soient Q de telles coordonnées; alors puisque $\dot{Q} = C\dot{q}$, on a

$$T=\frac{1}{2}\sum_{i=1}^{n}\dot{Q}_i^2,\quad U=\frac{1}{2}\sum_{i=1}^{n}\lambda_i Q_i^2. \tag{2}$$

Les nombres λ_i sont appelés *valeurs propres de la forme B relativement à A*.

Exercice. Montrer que les *valeurs propres de B relativement à A vérifient l'équation caractéristique*

$$\det|B-\lambda A| = 0, \tag{3}$$

dont toutes les racines sont, de ce fait, réelles (les matrices associées à A et B sont symétriques, $A > 0$).

B. Oscillations propres. Dans les coordonnées Q le système de Lagrange se décompose en n équations indépendantes

$$\ddot{Q}_i = -\lambda_i Q_i. \tag{4}$$

Donc on a démontré le

Théorème. *Un système effectuant des petites oscillations est le produit cartésien de n systèmes d'ordre* 1 *effectuant des petites oscillations.*

Pour chaque système d'ordre 1 il importe de distinguer trois cas.

Premier cas. $\lambda = \omega^2 > 0$; la solution est $Q = C_1 \cos \omega t + C_2 \sin \omega t$ (oscillations).

*) Au besoin on peut introduire une structure euclidienne en prenant la première forme pour carré scalaire, puis réduire la deuxième forme à ses axes principaux par une transformation orthogonale au sens de cette structure euclidienne.

D e u x i è m e c a s. $\lambda = 0$; la solution est $Q = C_1 + C_2t$ (équilibre indifférent).

T r o i s i è m e c a s. $\lambda = -k^2 < 0$; la solution est $Q = C_1 \operatorname{ch} kt + C_2 \operatorname{sh} kt$ (instabilité).

Corollaire. *Supposons que l'une des valeurs propres* (3) *est positive*: $\lambda = \omega^2 > 0$. *Alors le système* (1) *peut effectuer des oscillations périodiques de la forme*

$$q(t) = (C_1 \cos \omega t + C_2 \sin \omega t)\, \xi, \tag{5}$$

où ξ *est un vecteur propre associé à la valeur propre* λ (fig. 78)

$$B\xi = \lambda A \xi.$$

Ces oscillations sont le produit cartésien du mouvement de dimension 1 $Q_i = C_1 \cos \omega_i t + C_2 \sin \omega_i t$ par les mouvements triviaux $Q_j = 0$ $(j \neq i)$.

D é f i n i t i o n. Le mouvement périodique (5) est appelé *oscillations propres* du système (1) et le nombre ω, *fréquence propre*.

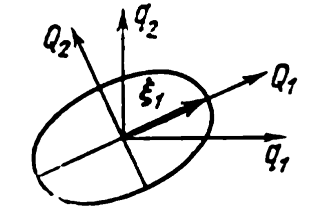


Fig. 78. Oscillations propres.

R e m a r q u e. Les oscillations et les fréquences propres sont également appelées principales ou normales. Des vecteurs propres sont également associés aux λ non positifs; pour simplifier nous appellerons les mouvements correspondants « oscillations propres » bien qu'ils ne soient pas périodiques; les « fréquences propres » correspondantes sont imaginaires.

E x e r c i c e. Montrer que le nombre d'oscillations propres (réelles) linéairement indépendantes est égal à l'indice positif d'inertie de l'énergie potentielle $^1/_2(Bq, q)$.

On peut maintenant formuler ce résultat sous la forme suivante:

Théorème. *Le système* (1) *possède n oscillations propres dont les directions sont deux à deux orthogonales au sens du produit scalaire défini par l'énergie cinétique A.*

En effet, le système de coordonnées Q en vertu de (2) est orthogonal au sens du carré scalaire (Aq, q).

C. Décomposition suivant les oscillations propres. Le théorème démontré entraîne le

Corollaire. *Toute petite oscillation est une somme d'oscillations propres.*

Une somme d'oscillations propres, d'une façon générale, n'est pas périodique (qu'on se rappelle les figures de Lissajous!).

Pour décomposer un mouvement en une somme d'oscillations propres il suffit de projeter les conditions initiales q, $\dot{q}$ sur les directions propres ξ_i et résoudre les problèmes de dimension 1 correspondants (4).

Donc les équations de Lagrange pour le système (1) peuvent se résoudre de la manière suivante. On cherche d'abord les oscillations propres sous la forme $q = e^{i\omega t}\xi$. En les portant ensuite dans les équations de Lagrange

$$\frac{d}{dt} A\dot{q} = Bq$$

on obtient

$$(B - \omega^2 A)\,\xi = 0.$$

De l'équation caractéristique (3) on déduit les n valeurs propres $\lambda_k = \omega_k^2$ auxquelles sont associés n vecteurs propres ξ_k deux à deux orthogonaux. Si $\lambda \neq 0$, la solution générale s'écrit

$$q(t) = \operatorname{Re} \sum_{k=1}^{n} C_k e^{i\omega_k t} \xi_k.$$

R e m a r q u e. Ce résultat est valable également lorsque certaines valeurs propres λ sont multiples.

Donc, *dans un système lagrangien, contrairement à un système général d'équations différentielles linéaires, les termes résonnants de la forme* $t \sin \omega t$, *etc. n'apparaissent pas même si certaines valeurs propres sont multiples.*

D. Exemples.

E x e m p l e 1. Considérons un système composé de deux pendules mathématiques identiques de longueurs $l_1 = l_2 = 1$, de masses $m_1 = m_2 = 1$ dans

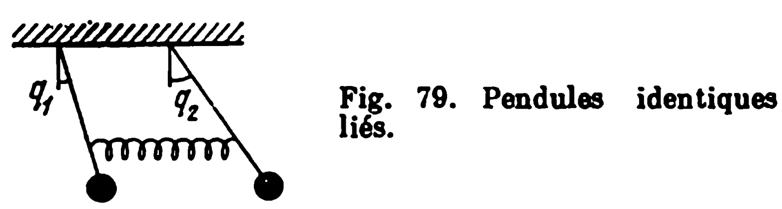


Fig. 79. Pendules identiques liés.

un champ de gravitation avec $g = 1$. Supposons que les pendules sont reliés par un ressort impondérable, dont la longueur est égale à la distance des points de suspension (fig. 79). Désignons par q_1, q_2 les angles d'écart des pendules. Alors pour de petites oscillations on a

$$T = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2), \quad U = \frac{1}{2}(q_1^2 + q_2^2) + \alpha (q_1 - q_2)^2,$$

où $\frac{\alpha}{2}(q_1 - q_2)^2$ est l'énergie potentielle d'élasticité du ressort. Posons:

$$Q_1 = \frac{q_1 + q_2}{\sqrt{2}}, \quad Q_2 = \frac{q_1 - q_2}{\sqrt{2}},$$

alors

$$q_1=\frac{Q_1+Q_2}{\sqrt{2}}, \quad q_2=\frac{Q_1-Q_2}{\sqrt{2}},$$

et les deux formes sont rapportées à leurs axes principaux:

$$T=\frac{1}{2}(\dot{Q}_1^2+\dot{Q}_2^2), \quad U=\frac{1}{2}(\omega_1^2Q_1^2+\omega_2^2Q_2^2),$$

où $\omega_1 = 1$, $\omega_2 = \sqrt{1+2\alpha}$ (fig. 80). Donc nous avons deux oscillations propres suivantes (fig. 81):

1. $Q_2 = 0$, i.e. $q_1 = q_2$: les deux pendules se déplacent en phase avec la fréquence initiale 1, le ressort n'est pas sollicité;

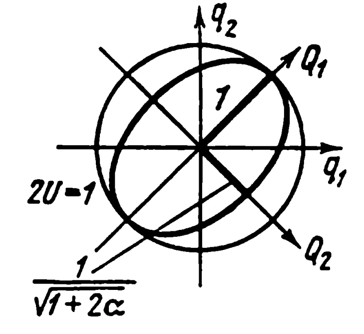


Fig. 80. Espace de configuration des pendules liés.

2. $Q_1 = 0$, i.e. $q_1 = -q_2$: les pendules sont en opposition de phase, leur fréquence $\omega_2 > 1$ s'accroît grâce à l'action du ressort.

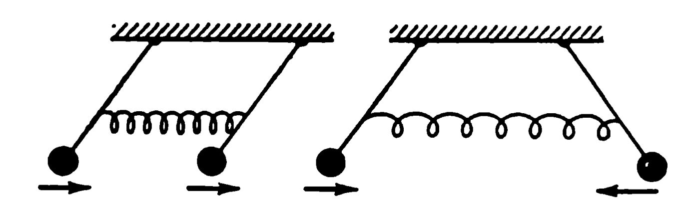


Fig. 81. Oscillations propres des pendules liés.

Supposons maintenant que le ressort est très faible: $\alpha \ll 1$. On assiste alors à un intéressant effet de *pompage d'énergie*.

E x e m p l e 2. Supposons que les pendules sont au repos à la date initiale et que l'un d'eux est animé de la vitesse $\dot{q}_1 = v$. Montrer qu'*au bout d'un temps T le premier pendule sera pratiquement immobile et toute l'énergie passera sur le second.*

Des conditions initiales il résulte que $Q_1(0) = Q_2(0) = 0$. Donc $Q_1 = c_1 \sin t$, $Q_2 = c_2 \sin \omega t$, $\omega = \sqrt{1+2\alpha} \approx 1+\alpha$ $(\alpha \ll 1)$. Or $\dot{Q}_1(0) = \dot{Q}_2(0) = \frac{v}{\sqrt{2}}$. Donc $c_1 = \frac{v}{\sqrt{2}}$, $c_2 = \frac{v}{\omega\sqrt{2}}$ et notre solution est

$$q_1=\frac{v}{2}\left(\sin t+\frac{1}{\omega}\sin \omega t\right), \quad q_2=\frac{v}{2}\left(\sin t-\frac{1}{\omega}\sin \omega t\right),$$

ou en négligeant le terme $v\left(1-\frac{1}{\omega}\right)$ sin ωt qui est petit avec α

$$q_1 \approx \frac{v}{2}(\sin t + \sin \omega t) = v \cos \varepsilon t \sin \omega' t,$$

$$q_2 \approx \frac{v}{2}(\sin t - \sin \omega t) = -v \cos \omega' t \sin \varepsilon t,$$

$$\varepsilon = \frac{\omega - 1}{2} = \frac{\alpha}{2}, \quad \omega' = \frac{\omega + 1}{2} \approx 1.$$

La grandeur $\varepsilon \approx \frac{\alpha}{2}$ est petite avec α, donc q_1 subit des oscillations de fréquence $\omega' \approx 1$ d'amplitude $v \cos \varepsilon t$ faiblement variable (fig. 82).

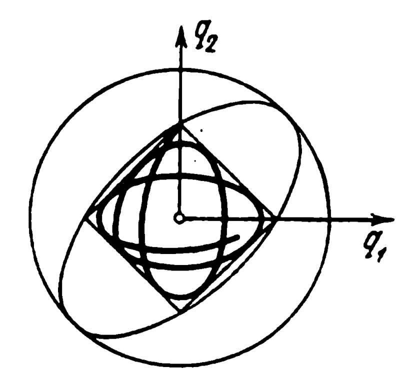


Fig. 82. Battements : trajectoire dans l'espace de configuration.

Au bout d'un temps $T = \frac{\pi}{2\varepsilon} \approx \frac{\pi}{\alpha}$ seul le deuxième pendule oscillera, au bout d'un temps $2T$, seul le premier oscillera et ainsi de suite (phénomène de battement) (fig. 83).

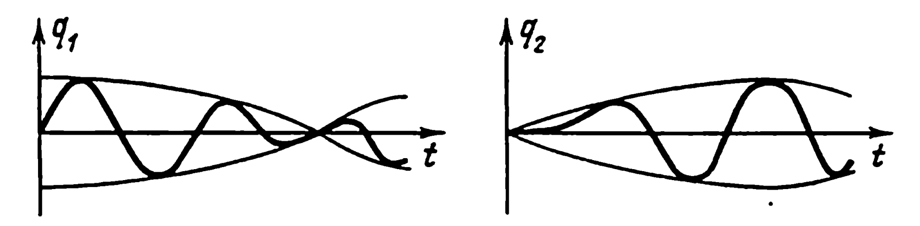


Fig. 83. Battements.

E x e m p l e 3. Etudier les *oscillations propres de deux pendules différents* ($m_1 \neq m_2$, $l_1 \neq l_2$, $g = 1$) *reliés par un ressort d'énergie* $^1/_2\alpha\,(q_1 - q_2)^2$ (fig. 84). *Comment se comportent les fréquences propres pour* $\alpha \to 0$ *et pour* $\alpha \to \infty$?

On a

$$T = {}^1/_2\,(m_1 l_1^2 \dot{q}_1^2 + m_2 l_2^2 \dot{q}_2^2),$$

$$U = m_1 l_1 \frac{q_1^2}{2} + m_2 l_2 \frac{q_2^2}{2} + \frac{\alpha}{2}(q_1 - q_2)^2.$$

Donc (fig. 85)

$$A = \begin{vmatrix} m_1 l_1^2 & 0 \\ 0 & m_2 l_2^2 \end{vmatrix},$$

$$B = \begin{vmatrix} m_1 l_1 + \alpha & -\alpha \\ -\alpha & m_2 l_2 + \alpha \end{vmatrix},$$

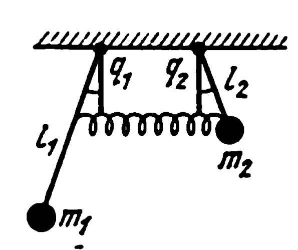


Fig. 84. Pendules liés.

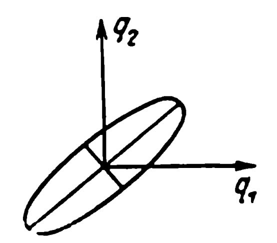


Fig. 85. Energie potentielle de pendules fortement liés.

et l'équation caractéristique s'écrit

$$\det|B-\lambda A| = \det\begin{vmatrix} m_1l_1+\alpha-\lambda m_1l_1^2 & -\alpha \\ -\alpha & m_2l_2+\alpha-\lambda m_2l_2^2 \end{vmatrix} = 0,$$

ou

$$a\lambda^2 - (b_0 + b_1\alpha)\lambda + (c_0 + c_1\alpha) = 0,$$

où

$$a = m_1m_2l_1^2l_2^2,$$

$$b_0 = m_1l_1m_2l_2(l_1 + l_2), \quad b_1 = m_1l_1^2 + m_2l_2^2,$$

$$c_0 = m_1m_2l_1l_2, \quad c_1 = m_1l_1 + m_2l_2.$$

C'est l'équation d'une hyperbole dans le plan (α, λ) (fig. 86). Lorsque $\alpha \to 0$ (ressort faible), les fréquences tendent vers les fréquences de pendules

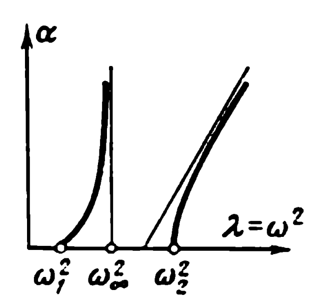


Fig. 86. Dépendance des fréquences propres par rapport à la rigidité du ressort.

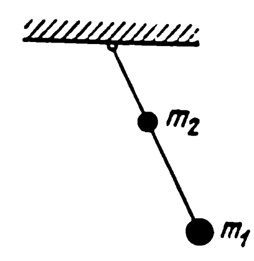


Fig. 87. Cas limite de pendules liés par un ressort infiniment rigide.

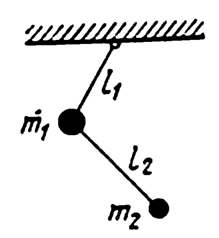
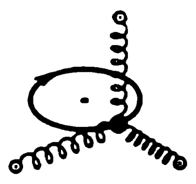


Fig. 88. Pendule double.


Fig. 89. Système dont l'ensemble des oscillations propres est infini.

libres $(\omega_{1,2}^2 = l_{1,2}^{-1})$; pour $\alpha \to \infty$ (un ressort très fort) l'une des fréquences tend vers ∞ et l'autre vers la fréquence propre ω_∞ du pendule double (fig. 87):

$$\omega_\infty^2 = \frac{m_1l_1 + m_2l_2}{m_1l_1^2 + m_2l_2^2}.$$

Exercice. Etudier les oscillations propres du pendule double plan (fig. 88).

Exercice. Déterminer la forme des trajectoires des petites oscillations d'un point matériel situé au centre d'un triangle équilatéral et lié par des ressorts identiques aux sommets (fig. 89).

Solution. Le système est invariant par une rotation de 120°. Donc toutes les directions sont propres et les deux fréquences propres sont égales, de sorte que $U = \frac{1}{2}\omega^2(x^2 + y^2)$. Donc les trajectoires sont des ellipses (voir fig. 20).

§ 24. Comportement des fréquences propres

On démontre ici les théorèmes de Rayleigh-Courant-Fisher sur le comportement des fréquences propres d'un système à l'accroissement de la rigidité et en présence de liaisons.

A. Comportement des fréquences propres à la variation de la rigidité.

Considérons les petites oscillations d'un système d'énergies cinétique et potentielle respectivement

$$T = \frac{1}{2}(A\dot{q}, \dot{q}) > 0, \quad U = \frac{1}{2}(Bq, q) > 0 \quad \forall q, \dot{q} \neq 0.$$

Définition. Un système de même énergie cinétique T et d'énergie potentielle U' est *plus rigide* si

$$U' = \frac{1}{2}(B'q, q) \geqslant \frac{1}{2}(Bq, q) = U \ \forall q.$$

Nous allons voir comment varient les fréquences propres lorsque la rigidité du système croît.

Exercice. Traiter le cas unidimensionnel.

Théorème 1. *Lorsque la rigidité croît, toutes les fréquences propres croissent, i.e. si* $\omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n$ *sont les fréquences propres du système moins rigide et* $\omega_1' \leqslant \omega_2' \leqslant \ldots \leqslant \omega_n'$ *les fréquences du système plus rigide, alors* $\omega_1 \leqslant \omega_1'$; $\omega_2 \leqslant \omega_2'$; $\ldots$; $\omega_n \leqslant \omega_n'$.

Ce théorème admet une traduction géométrique simple. Sans nuire à la généralité on peut supposer que $A = E$, i.e. nous considérons une structure euclidienne définie par l'énergie cinétique $T = \frac{1}{2}(\dot{q}, \dot{q})$. Associons à chaque système les ellipsoïdes E: $(Bq, q) = 1$, E': $(B'q, q) = 1$ respectivement. Il est évident que

Lemme 1. *Si le système U' est plus rigide que le système U, alors l'ellipsoïde* E' *qui lui est associé est contenu dans* E.

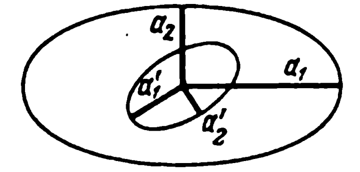
Est aussi évident le

Lemme 2. *Les demi-axes principaux de l'ellipsoïde* E *sont inverses des fréquences* ω_i:

$$\omega_i = \frac{1}{a_i}.$$

Donc le théorème 1 équivaut à la proposition géométrique suivante (fig. 90):

Fig. 90. Les demi-axes de l'ellipsoïde intérieur sont plus petits.



Théorème 2. *Si un ellipsoïde* E *de demi-axes* $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$ *contient un ellipsoïde* E_1 *de demi-axes* $a'_1 \geqslant a'_2 \geqslant \ldots \geqslant a'_n$ *et de même centre, alors les demi-axes de l'ellipsoïde* E_1 *sont plus petits*:

$$a_1 \geqslant a'_1, \quad a_2 \geqslant a'_2, \ \ldots, \ a_n \geqslant a'_n.$$

Exemple. L'accroissement de la rigidité α du ressort liant les deux pendules de l'exemple 3, § 23 entraîne celle de l'énergie potentielle et en vertu du théorème 1 celle des fréquences propres: $\frac{d\omega_i}{d\alpha} > 0$. Traitons maintenant le cas où la rigidité du ressort tend vers l'infini: $\alpha \to \infty$. A la limite, les pendules sont rigidement liés et l'on obtient un système à un degré de liberté; la fréquence propre limite est $\omega_1 < \omega_\infty < \omega_2$.

B. Comportement des fréquences propres en présence de liaisons. Reprenons le système général à n degrés de liberté et soient $T = \frac{1}{2}(\dot{q}, \dot{q})$, $U = \frac{1}{2}(Bq, q)$, $q \in \mathbf{R}^n$ les énergies cinétique et potentielle de ce système en petites oscillations.

Soit $\mathbf{R}^{n-1} \subset \mathbf{R}^n$ un sous-espace de dimension $n-1$ de $\mathbf{R}^n$ (fig. 91). Considérons un système à $n-1$ degrés de liberté ($q \in \mathbf{R}^{n-1}$)

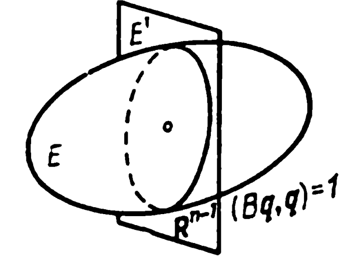


Fig. 91. Liaison linéaire.

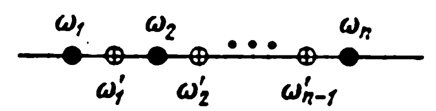


Fig. 92. Séparation des fréquences.

dont les énergies cinétique et potentielle sont égales aux restrictions de T et de U à $\mathbf{R}^{n-1}$. On dit que ce système est déduit du système initial *par assujetissement à une liaison linéaire.*

Supposons que le système initial possède n fréquences propres $\omega_1 \leqslant \omega_2 \leqslant \ldots \leqslant \omega_n$ et le système soumis à une liaison $n-1$

fréquences propres

$$\omega_1' \leqslant \omega_2' \leqslant \ldots \leqslant \omega_{n-1}'.$$

Théorème 3. *Les fréquences propres du système à une liaison séparent les fréquences propres du système initial* (fig. 92):

$$\omega_1 \leqslant \omega_1' \leqslant \omega_2 \leqslant \omega_2' \leqslant \ldots \leqslant \omega_{n-1} \leqslant \omega_{n-1}' \leqslant \omega_n.$$

En vertu du lemme 2 ce théorème est équivalent à la proposition géométrique suivante.

Théorème 4. *Considérons une section* E′ *d'un ellipsoïde* $E = \{q : (Bq, q) = 1\}$ *de demi-axes* $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$ *par l'hyper-*

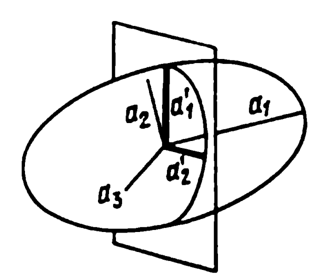


Fig. 93. Les demi-axes de la section séparent les demi-axes de l'ellipsoïde.

plan R^{n-1}. *Alors les demi-axes de l'ellipsoïde* E′ *de dimension* $n - 1$ *séparent les demi-axes de l'ellipsoïde* E (fig. 93):

$$a_1 \geqslant a_1' \geqslant a_2 \geqslant a_2' \geqslant \ldots \geqslant a_{n-1} \geqslant a_{n-1}' \geqslant a_n.$$

C. Propriétés extrémales des valeurs propres.

Théorème 5. *Toute section d'un ellipsoïde* E *de demi-axes* $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_k$ *par un sous-espace* R^k *de dimension* k *a son demi petit axe plus petit ou égal à* a_k:

$$a_k = \max_{\{R^k\}} \min_{x \in R^k \cap E} \| x \|$$

(*la borne supérieure est atteinte sur le sous-espace engendré par les demi-axes* $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_k$).

D é m o n s t r a t i o n *). Considérons le sous-espace R^{n-k+1} engendré par les axes $a_k \geqslant a_{k+1} \geqslant \ldots \geqslant a_n$. Il est de dimension $n - k + 1$. Il coupera donc R^k. Soit x un point d'intersection situé sur l'ellipsoïde E. On a $\| x \| \leqslant a_k$, puisque $x \in R^{n-k+1}$. Comme $\| x \|$ n'est pas inférieur à la longueur du demi petit axe de l'ellipsoïde $E \cap R^k$, ce demi petit axe n'est pas supérieur à a_k, c.q.f.d.

D é m o n s t r a t i o n d u t h é o r è m e 2. Le demi petit axe de chaque section (de dimension k) de l'ellipsoïde intérieur $R^k \cap E'$ n'est pas supérieur au demi petit axe de la section $R^k \cap E$. En vertu du théorème 5

$$a_k' = \max_{\{R^k\}} \min_{x \in R^k \cap E'} \| x \| \leqslant \max_{\{R^k\}} \min_{x \in R^k \cap E} \| x \| = a_k, \text{ c.q.f.d.}$$

*) Il est utile de se représenter le cas $n = 3$, $k = 2$.

D é m o n s t r a t i o n d u t h é o r è m e 4. L'inégalité $a'_k \leqslant a_k$ découle du théorème 5, puisque dans le calcul de a_k le maximum est pris sur un ensemble plus vaste. Pour prouver l'inégalité $a'_k \geqslant a_{k+1}$ coupons $\mathbf{R}^{n-1}$ par un sous-espace quelconque $\mathbf{R}^{k+1}$ de dimension $k+1$. La section aura une dimension non inférieure à k. Le demi petit axe de l'ellipsoïde $E' \cap \mathbf{R}^{k+1}$ sera non inférieur au demi petit axe de $E \cap \mathbf{R}^{k+1}$. En vertu du théorème 5

$$a'_k = \max_{\{\mathbf{R}^k \subset \mathbf{R}^{n-1}\}} \min_{x \in \mathbf{R}^k \cap E'} \|x\| \geqslant \max_{\{\mathbf{R}^{k+1} \subset \mathbf{R}^n\}} \min_{x \in \mathbf{R}^{k+1} \cap E'} \|x\| \geqslant$$
$$\geqslant \max_{\{\mathbf{R}^{k+1} \subset \mathbf{R}^n\}} \min_{x \in \mathbf{R}^{k+1} \cap E} \|x\| = a_{k+1}, \text{ c.q.f.d.}$$

Les théorèmes 1 et 3 découlent immédiatement des théorèmes que nous venons de démontrer.

E x e r c i c e. Montrer que si, laissant invariante l'énergie potentielle d'un système, on augmente l'énergie cinétique (par exemple, on conserve les ressorts et l'on augmente les masses), chaque fréquence propre diminuera.

E x e r c i c e. Montrer que tous les demi-axes d'un ellipsoïde situé dans un sous-espace d'un espace euclidien diminuent par projection orthogonale sur un autre sous-espace.

E x e r c i c e. Soit $A(\varepsilon)$ une forme quadratique sur l'espace euclidien $\mathbf{R}^n$ continûment différentiable par rapport au paramètre ε. Montrer que toute valeur propre est continûment différentiable par rapport à ε et trouver ses dérivées.

Réponse. Soient $\lambda_1, \ldots, \lambda_k$ les valeurs propres de $A(0)$. A chaque valeur propre λ_i de multiplicité ν_i est associé un sous-espace $\mathbf{R}^{\nu_i}$. *Les dérivées des valeurs propres de $A(\varepsilon)$ sont égales en 0 aux valeurs propres de la restriction de la forme* $B = \left.\dfrac{dA}{d\varepsilon}\right|_{\varepsilon=0}$ à $\mathbf{R}^{\nu_i}$.

En particulier, *si toutes les valeurs propres de $A(0)$ sont simples, leurs dérivées sont égales aux éléments diagonaux de la matrice B dans une base propre de $A(0)$.*

De cet exercice l'on déduit donc que si la forme s'accroît, ses valeurs propres s'accroissent également. On obtient ainsi une nouvelle démonstration des théorèmes 1 et 2.

E x e r c i c e. Etudier l'influence d'une fêlure sur la hauteur du son d'une cloche.

§ 25. Résonance paramétrique

Si les paramètres d'un système sont périodiques, la position d'équilibre peut devenir instable même si elle est stable pour chaque valeur fixe d'un paramètre. C'est d'ailleurs cette instabilité qui est à l'origine des mouvements de la balançoire.

A. Systèmes dynamiques à paramètres périodiques.

E x e m p l e 1. Une *balançoire* dont la longueur $l(t)$ du pendule mathématique équivalent varie périodiquement en fonction du temps: $l(t+T) = l(t)$ (fig. 94).

E x e m p l e 2. *Un pendule dans un champ de force de pesanteur périodique* (par exemple la Lune) *obéit à l'équation de Hill*

$$\ddot{q} = -\omega^2 (t) q, \quad \omega (t + T) = \omega (t). \tag{1}$$

E x e m p l e 3. *Un pendule dont le point d'attache est animé d'oscillations périodiques verticales est également décrit par une équation de la forme* (1).

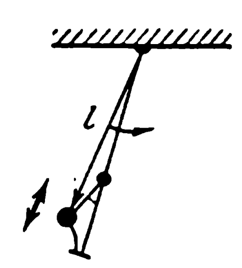


Fig. 94. Balançoire.

Dans les systèmes à paramètres périodiques, les seconds membres des équations de mouvement sont des fonctions périodiques de t. Les équations de mouvement peuvent être écrites sous forme d'un système d'équations différentielles du premier ordre

$$\dot{x} = f (x, t), \quad f (x, t + T) = f (x, t), \quad x \in \mathbf{R}^n, \tag{2}$$

à seconds membres périodiques. Par exemple, l'équation (1) peut se mettre sous la forme du système

$$\left.\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\omega^2 x_1 \end{aligned}\right\} \omega (t + T) = \omega (t). \tag{3}$$

B. Application pendant une période. Rappelons les propriétés générales des systèmes (2). Désignons par g^t: $\mathbf{R}^n \to \mathbf{R}^n$ une application qui à $x \in \mathbf{R}^n$ associe la valeur $g^t x = \varphi (t)$ prise à la date t par la solution φ du système (2) qui vérifie la condition initiale $\varphi (0) = x$ (fig. 95).

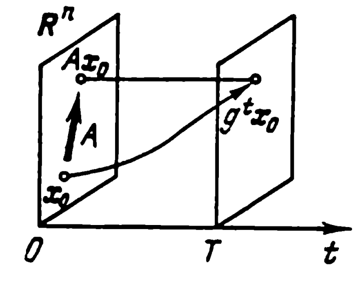


Fig. 95. Application pendant une période.

Les applications g^t ne forment pas de groupe : d'une façon générale

$$g^{t+s} \neq g^t g^s \neq g^s g^t.$$

Exercice. Démontrer qu'*une condition nécessaire et suffisante pour que* $\{g^t\}$ *soit groupe est que les seconds membres* f *ne dépendent pas de* t.

Exercice. *Montrer que si* T *est la période de* f, *alors* $g^{T+s} = g^s \cdot g^T$ *et, en particulier,* $g^{nT} = (g^T)^n$, *de sorte que les applications* g^{nT} *forment un groupe* (n est entier).

L'application $g^T : \mathbf{R}^n \to \mathbf{R}^n$ jouera un rôle important dans la suite ; nous l'appellerons *application pendant une période* et la noterons

$$A : \mathbf{R}^n \to \mathbf{R}^n, \quad Ax(0) = x(T).$$

Exemple. Pour les systèmes

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1, \end{cases} \qquad \begin{cases} \dot{x}_1 = x_1, \\ \dot{x}_2 = -x_2 \end{cases}$$

que nous pouvons considérer comme périodiques et de période quelconque T, l'application A est respectivement une rotation et une rotation hyperbolique (fig. 96).

Théorème. 1) *Un point* x_0 *est point fixe de l'application* A ($Ax_0 = x_0$) *si et seulement si la solution réalisant la condition initiale* $x(0) = x_0$ *est périodique et de période* T.

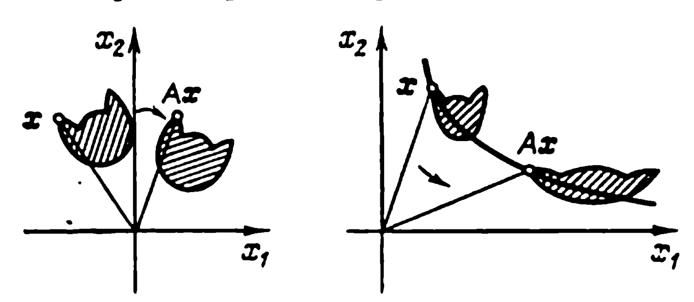


Fig. 96. Rotation simple et rotation hyperbolique.

2) *Une solution périodique* $x(t)$ *est stable au sens de Liapounov (resp. asymptotiquement stable) si et seulement si le point fixe* x_0 *de l'application* A *est stable au sens de Liapounov (resp. asymptotiquement stable)* *).

3) *Si le système* (2) *est linéaire, i.e.* $f(x, t) = f(t)\,x$ *est une fonction linéaire de* x, *alors l'application* A *est linéaire.*

4) *Si le système* (2) *est hamiltonien, alors l'application* A *conserve le volume* $\det A = 1$.

---

*) Le point fixe x_0 de l'application A est *stable au sens de Liapounov* (resp. *asymptotiquement stable*) si $\forall \varepsilon > 0$, $\exists \delta > 0$ tel que $|x - x_0| < \delta$ entraîne $|A^n x - A^n x_0| < \varepsilon$ pour tous les n $(0 < n < \infty)$ à la fois (resp. $|A^n x - A^n x_0| \to 0$ pour $n \to \infty$).

D é m o n s t r a t i o n. 1) et 2) découlent de la relation $g^{T+s} = g^s A$. 3) découle du fait que la somme des solutions d'un système linéaire est de nouveau solution. 4) résulte du théorème de Liouville.

Appliquons le théorème démontré à l'application A du plan des phases $\{(x_1, x_2)\}$ sur lui-même, correspondant à l'équation (1) et au système (3). Comme le système (3) est linéaire et hamiltonien $\left(H = \frac{x_1^2}{2} + \omega^2 \frac{x_2^2}{2}\right)$, il vient le

Corollaire. *L'application A est linéaire et conserve l'aire* ($\det A = 1$). *Pour que la solution triviale de l'équation* (1) *soit stable il faut et il suffit que l'application A le soit.*

E x e r c i c e. Montrer que la rotation du plan est une application stable et la rotation hyperbolique une application instable.

C. Applications linéaires du plan sur lui-même conservant l'aire.

Théorème. *Soit A la matrice d'une application linéaire du plan sur lui-même conservant l'aire* ($\det A = 1$). *L'application A est stable si* $|\operatorname{tr} A| < 2$ *et instable si* $|\operatorname{tr} A| > 2$ ($\operatorname{tr} A = a_{11} + a_{22}$).

D é m o n s t r a t i o n. Soient λ_1, λ_2 les valeurs propres de A. Elles vérifient l'équation caractéristique $\lambda^2 - \operatorname{tr} A\lambda + 1 = 0$ de coefficients réels $\lambda_1 + \lambda_2 = \operatorname{tr} A$, $\lambda_1 \cdot \lambda_2 = \det A = 1$.

Les racines λ_1, λ_2 de cette équation réelle du second degré sont réelles pour $|\operatorname{tr} A| > 2$ et conjuguées complexes pour $|\operatorname{tr} A| < 2$.

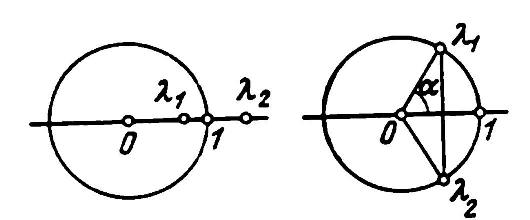


Fig. 97. Valeurs propres de l'application A.

Dans le premier cas l'une des valeurs propres est plus grande et l'autre plus petite que l'unité en module; l'application A est une rotation hyperbolique, donc elle est instable (fig. 97).

Dans le deuxième cas les deux valeurs propres sont situées sur le cercle unité (fig. 97):

$$1 = \lambda_1 \cdot \lambda_2 = \lambda_1 \cdot \bar{\lambda}_1 = |\lambda_1|^2.$$

L'application A équivaut à une rotation d'angle α (où $\lambda_{1,2} = e^{\pm i\alpha}$) i.e. se réduit à une rotation par un choix convenable des coordonnées sur le plan. Donc elle est stable, c.q.f.d.

Par conséquent la question de savoir si la solution triviale de l'équation (1) est stable se ramène au calcul de la trace de la matrice A. Malheureusement on ne peut expliciter cette trace que dans des cas spéciaux. La plupart du temps on la détermine approximativement

par intégration numérique de l'équation sur l'intervalle $0 \leqslant t \leqslant T$. Dans le cas important où $\omega(t)$ est proche d'une constante, on fait appel à des raisonnements généraux simples.

D. Stabilité forte.

D é f i n i t i o n. La solution triviale d'un système linéaire hamiltonien est *fortement stable* si elle est stable et si est également

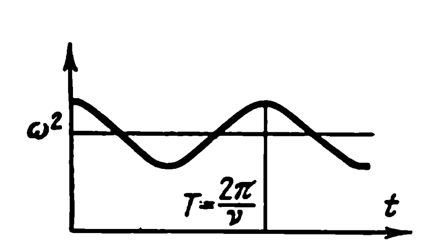


Fig. 98. Fréquence instantanée comme fonction du temps.

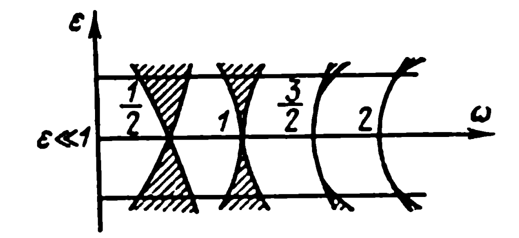


Fig. 99. Zones de résonance paramétrique.

stable la solution triviale d'un système linéaire hamiltonien suffisamment proche *).

Les deux théorèmes précédents entraînent le

Corollaire. *Si* $|\operatorname{tr} A| < 2$, *alors la solution triviale est fortement stable.*

En effet, si $|\operatorname{tr} A| < 2$, l'application A' correspondant à un système suffisamment proche est également telle que $|\operatorname{tr} A'| < 2$, c.q.f.d.

Appliquons ceci à un système à coefficients presque constants (variant peu). Soit par exemple l'équation

$$\ddot{x} = -\omega^2 (1 + \varepsilon a(t))\, x, \quad \varepsilon \ll 1, \tag{4}$$

où $a(t + 2\pi) = a(t)$ avec $a(t) = \cos t$ par exemple (fig. 98). (Un pendule dont la fréquence oscille près de ω, de faible amplitude et de période 2π.) **)

Chaque système (4) sera représenté par un point du plan engendré par les paramètres ε, $\omega > 0$. Il est évident que les systèmes stables ($|\operatorname{tr} A| < 2$) forment sur le plan (ω, ε) un ensemble ouvert de même d'ailleurs que les systèmes instables ($|\operatorname{tr} A| > 2$) (fig. 99).

La frontière de stabilité est définie par l'équation $|\operatorname{tr} A| = 2$.

---

*) La distance de deux systèmes linéaires à coefficients périodiques $\dot{x} = B_1(t)x$, $\dot{x} = B_2(t)x$ se définit comme le maximum de la distance des opérateurs $B_1(t)$ et $B_2(t)$ par rapport à t.

**) Lorsque $a(t) = \cos t$, l'équation (4) est appelée *équation de Mathieu*.

Théorème. *Tous les points de l'axe ω à l'exception des entiers et demi-entiers* $\omega = \frac{k}{2}$, $k = 0, 1, 2, \ldots$, *correspondent à des systèmes fortement stables* (4).

Donc l'ensemble des systèmes instables est susceptible d'approcher l'axe ω seulement aux points $\omega = k/2$. En d'autres termes, on ne peut faire balancer une balançoire, par de petites variations périodiques de sa longueur, que si une période de variation de la longueur est proche d'un nombre entier de demi-périodes des oscillations propres, un résultat que tout le monde connaît empiriquement.

La démonstration du théorème formulé est basée sur le fait que pour $\varepsilon = 0$ l'équation (4) possède des coefficients constants et se résout explicitement.

Exercice. *Calculer pour le système* (4) *avec* $\varepsilon = 0$ *la matrice de l'application* A *pendant la période* $T = 2\pi$ *dans la base* $x, \dot{x}$.

Solution. La solution générale est

$$x = c_1 \cos \omega t + c_2 \sin \omega t.$$

La solution particulière vérifiant la condition initiale $x = 1$, $\dot{x} = 0$ est

$$x = \cos \omega t, \quad \dot{x} = -\omega \sin \omega t.$$

La solution particulière réalisant la condition initiale $x = 0$, $\dot{x} = 1$ est

$$x = \frac{1}{\omega} \sin \omega t, \quad \dot{x} = \cos \omega t.$$

Réponse.

$$A = \begin{vmatrix} \cos 2\pi\omega & \frac{1}{\omega} \sin 2\pi\omega \\ -\omega \sin 2\pi\omega & \cos 2\pi\omega \end{vmatrix}.$$

Donc $|\operatorname{tr} A| = |2 \cos 2\omega\pi| < 2$ si $\omega \neq \frac{k}{2}$, $k = 0, 1, \ldots$, et le théorème découle du corollaire précédent.

Une analyse plus approfondie *) montre que d'une façon générale (et pour $a(t) = \cos t$) au voisinage des points $\omega = \frac{k}{2}$, $k = 1, 2, \ldots$, le domaine d'instabilité (hachuré sur la fig. 99) s'approche effectivement de l'axe ω.

Donc lorsque $\omega \approx \frac{k}{2}$, $k = 1, 2, \ldots$, la position inférieure d'équilibre d'une balançoire parfaite (4) est instable et elle se balance pour une variation périodique de la longueur aussi petite que l'on

*) Voir par exemple l'exercice de la page 123.

veut. Ce phénomène s'appelle *résonance paramétrique*. La résonance paramétrique a ceci de particulier qu'elle se manifeste le plus fortement lorsque la fréquence de variation des paramètres ν ($\nu = 1$ dans l'équation (4)) est le double de la fréquence propre ω.

R e m a r q u e. Théoriquement la résonance paramétrique devrait s'observer pour une infinité de valeurs $\omega/\nu \approx k/2$, $k = 1, 2, \ldots$ En pratique elle ne se manifeste que lorsque $k = 1$ ou 2, rarement à 3. C'est que

a) pour les grands k le domaine d'instabilité s'approche de l'axe ω en affectant la forme d'une langue étroite et pour la fréquence de résonance ω on obtient des limites très strictes ($\sim \varepsilon^k$ pour la fonction différentiable $a(t)$ de (4));

b) l'instabilité est faiblement exprimée pour les grands k, puisque $|\operatorname{tr} A| - 2$ n'est pas élevé et les valeurs propres sont proches de l'unité;

c) quelque petit que soit le frottement il existe une valeur minimale $\varepsilon_k \neq 0$ de l'amplitude telle que pour les valeurs $\varepsilon < \varepsilon_k$ il n'y a pas

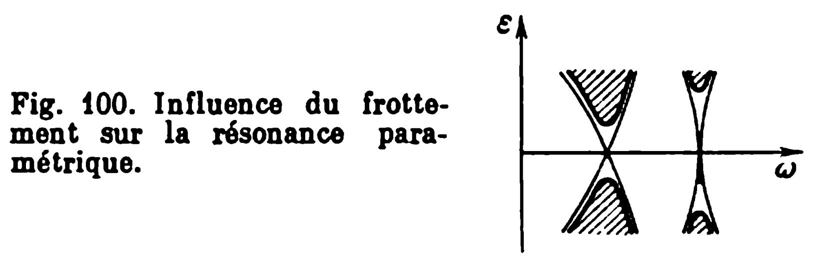


Fig. 100. Influence du frottement sur la résonance paramétrique.

de résonance paramétrique, i.e. les oscillations s'amortissent. ε_k croît très vite avec k (fig. 100).

Remarquons aussi que, s'agissant des équations (4), la grandeur x croît indéfiniment dans le cas instable.

Dans les systèmes réels les oscillations atteignent une amplitude finie, car pour les grands x l'équation linéarisée (4) n'est plus valable et il faut tenir compte des effets non linéaires.

E x e r c i c e. *Déterminer la forme des domaines de stabilité sur le plan* (ε, ω) *du système décrit par l'équation*

$$\ddot{x} = -f^2(t)\,x, \quad f(t) = \begin{cases} \omega+\varepsilon, & 0 < t < \pi, \\ \omega-\varepsilon, & \pi < t < 2\pi, \end{cases} \quad \varepsilon \ll 1,$$
$$f(t + 2\pi) = f(t).$$

S o l u t i o n. De l'exercice précédent il vient $A = A_2 A_1$, où

$$A_k = \begin{vmatrix} c_k & \dfrac{1}{\omega_k} s_k \\ -\omega_k s_k & c_k \end{vmatrix}, \quad \begin{array}{l} c_k = \cos \pi\omega_k, \\ s_k = \sin \pi\omega_k, \\ \omega_{1,2} = \omega \pm \varepsilon. \end{array}$$

Donc la frontière du domaine de stabilité a pour équation

$$|\operatorname{tr} A| = \left| 2c_1c_2 - \left(\frac{\omega_1}{\omega_2} + \frac{\omega_2}{\omega_1}\right) s_1 s_2 \right| = 2. \qquad (5)$$

Comme $\varepsilon \ll 1$, on a $\frac{\omega_1}{\omega_2}=\frac{\omega+\varepsilon}{\omega-\varepsilon} \approx 1$. Introduisons la notation

$$\frac{\omega_1}{\omega_2}+\frac{\omega_2}{\omega_1}=2(1+\Delta).$$

D'où il vient $\Delta = \frac{2\varepsilon^2}{\omega^2}+O(\varepsilon^4) \ll 1$. Compte tenu des expressions

$$2c_1c_2 = \cos 2\pi\varepsilon + \cos 2\pi\omega, \quad 2s_1s_2 = \cos 2\pi\varepsilon - \cos 2\pi\omega,$$

l'équation (5) s'écrit

$$-\Delta \cos 2\pi\varepsilon + (2+\Delta)\cos 2\pi\omega = \pm 2$$

ou

$$\cos 2\pi\omega = \frac{2+\Delta\cos 2\pi\varepsilon}{2+\Delta}, \tag{6_1}$$

$$\cos 2\pi\omega = \frac{-2+\Delta\cos 2\pi\varepsilon}{2+\Delta}. \tag{6_2}$$

Dans le premier cas $\cos 2\pi\omega \approx 1$. On posera donc

$$\omega = k+a, \quad |a| \ll 1; \quad \cos 2\pi\omega = \cos 2\pi a = 1 - 2\pi^2a^2 + O(a^4).$$

Transcrivons l'équation (6_1) sous la forme

$$\cos 2\pi\omega = 1-\frac{\Delta}{2+\Delta}(1-\cos 2\pi\varepsilon)$$

ou

$$2\pi^2a^2 + O(a^4) = \Delta\pi^2\varepsilon^2 + O(\varepsilon^4).$$

Compte tenu de $\Delta=\frac{2\varepsilon^2}{\omega^2}+O(\varepsilon^4)$ il vient

$$a=\pm\frac{\varepsilon^2}{\omega^2}+o(\varepsilon^2), \quad \text{i.e.} \quad \omega=k\pm\frac{\varepsilon^2}{k^2}+o(\varepsilon^2).$$

L'équation (6_2) se résout de façon analogue; on obtient finalement

$$\omega=k+\frac{1}{2}\pm\frac{\varepsilon}{\pi\left(k+\frac{1}{2}\right)}+o(\varepsilon).$$

Donc la réponse est de la forme représentée sur la figure 101.

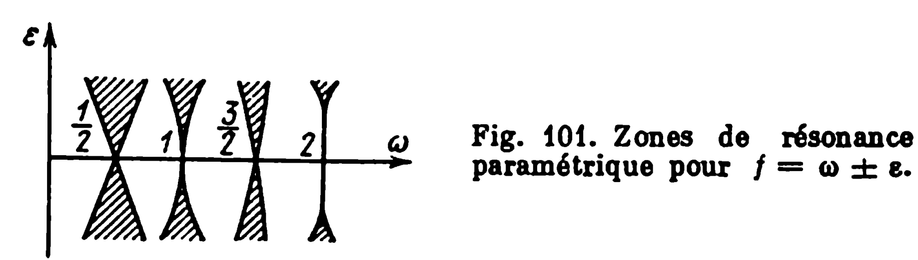


Fig. 101. Zones de résonance paramétrique pour $f = \omega \pm \varepsilon$.

E. Stabilité d'un pendule renversé dont le point d'attache oscille verticalement.

E x e r c i c e. *Est-ce que la position supérieure d'équilibre d'un pendule, d'ordinaire instable, peut devenir stable si son point d'attache oscille verticalement* (fig. 102)?

Soient l la longueur du pendule; $a \ll l$ l'amplitude des oscillations du point d'attache, 2τ la période du point d'attache, en outre pendant chaque demi-période l'accélération du point d'attache est

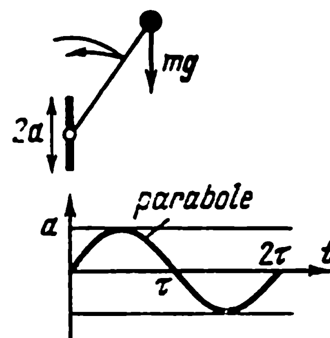



Fig. 102. Pendule renversé dont le point d'attache oscille.

constante et égale à $\pm c$ $\left(\text{alors } c = \frac{8a}{\tau^2}\right)$. Il se trouve que si les oscillations du point d'attache sont assez rapides ($\tau \ll 1$), la position supérieure d'équilibre devient stable.

S o l u t i o n. L'équation du mouvement peut se mettre sous la forme $\ddot{x} = (\omega^2 \pm d^2)\, x$ (le signe change au bout d'un temps τ), où $\omega^2 = g/l$, $d^2 = c/l$. Si les oscillations du point d'attache sont assez rapides, alors $d^2 > \omega^2$ $\left(d^2 = \frac{8a}{l\tau^2}\right)$.

Comme dans l'exercice précédent, $A = A_2 A_1$, où

$$A_1 = \begin{vmatrix} \operatorname{ch} k\tau & \frac{1}{k}\operatorname{sh} k\tau \\ k \operatorname{sh} k\tau & \operatorname{ch} k\tau \end{vmatrix}, \quad A_2 = \begin{vmatrix} \cos \Omega\tau & \frac{1}{\Omega}\sin \Omega\tau \\ -\Omega \sin \Omega\tau & \cos \Omega\tau \end{vmatrix}.$$

$$k^2 = d^2 + \omega^2, \qquad \Omega^2 = d^2 - \omega^2.$$

La condition de stabilité $|\operatorname{tr} A| < 2$ sera donc

$$\left| 2 \operatorname{ch} k\tau \cos \omega\tau + \left(\frac{k}{\Omega} - \frac{\Omega}{k}\right) \operatorname{sh} k\tau \sin \omega\tau \right| < 2. \tag{7}$$

Montrons que cette condition est réalisée pour des oscillations assez rapides du point d'attache, i.e. pour $c \gg g$. Introduisons les variables adimensionnées ε, μ:

$$\frac{a}{l} = \varepsilon^2 \ll 1, \quad \frac{g}{c} = \mu^2 \ll 1.$$

Alors

$$k\tau = 2\sqrt{2}\,\varepsilon \sqrt{1+\mu^2}, \quad \Omega\tau = 2\sqrt{2}\,\varepsilon\sqrt{1-\mu^2},$$

$$\frac{k}{\Omega} - \frac{\Omega}{k} = \sqrt{\frac{1+\mu^2}{1-\mu^2}} - \sqrt{\frac{1-\mu^2}{1+\mu^2}} = 2\mu^2 + O(\mu^4).$$

Donc pour ε et μ petits sont valables à $o(\varepsilon^4 + \mu^4)$ près les développements

$$\operatorname{ch} k\tau = 1 + 4\varepsilon^2 (1+\mu^2) + \frac{2}{3}\varepsilon^4 + \ldots, \quad \cos \Omega\tau = 1 - 4\varepsilon^2(1-\mu^2) + \frac{8}{3}\varepsilon^4 + \ldots,$$

$$\left(\frac{k}{\Omega} - \frac{\Omega}{k}\right) \operatorname{sh} k\tau \sin \Omega\tau = 16\varepsilon^2\mu^2 + \ldots$$

Et la condition de stabilité (7) s'écrit

$$2\left(1-16\varepsilon^4+\frac{16}{3}\varepsilon^4+8\varepsilon^2\mu^2+\ldots\right)+16\varepsilon^2\mu^2<2,$$

i.e. si l'on néglige les infiniment petits d'ordre supérieur, on a $\frac{2}{3}16\varepsilon^4 \geqslant 32\mu^2\varepsilon^2$ ou $\mu < \frac{\varepsilon}{\sqrt{3}}$ ou encore $\frac{g}{c} < \frac{a}{3l}$. Cette condition peut encore s'écrire

$$N > \sqrt{\frac{3}{32}}\,\omega\,\frac{l}{a} \approx 0{,}3\omega\frac{l}{a}\left(\sqrt{\frac{3}{32}} \approx 0{,}31\right),$$

où $N = \frac{1}{2\tau}$ est la fréquence des oscillations du point d'attache. Si par exemple la longueur du pendule est $l = 20$ cm et l'amplitude des oscillations du point d'attache $a = 1$ cm, alors

$$N > 0{,}31\sqrt{\frac{980}{20}}\cdot 20 \approx 43.$$

La position supérieure d'équilibre est stable si la fréquence des oscillations du point d'attache est par exemple supérieure à 50.

CHAPITRE 6

LE SOLIDE

Dans ce chapitre on se propose d'étudier en détail des problèmes de mécanique d'un type assez particulier. Ces problèmes sont traités traditionnellement dans les cours de mécanique classique pour deux raisons : d'abord parce qu'ils ont été résolus par Euler et Lagrange et ensuite parce que nous vivons dans un espace euclidien de dimension trois et la plupart des systèmes mécaniques à nombre fini de degrés de liberté que nous rencontrons sont des systèmes de solides.

§ 26. Mouvement par rapport à un système de coordonnées mobile

Dans ce paragraphe on définit la vitesse angulaire.

A. Systèmes de coordonnées mobiles. Considérons un système lagrangien décrit, dans les coordonnées q, t, par le lagrangien $L(q, \dot{q}, t)$. Il y a souvent intérêt à passer à un système de coordonnées mobile $Q = Q(q, t)$.

Pour écrire les équations de mouvement par rapport à un système mobile il suffit d'exprimer le lagrangien dans les nouvelles coordonnées.

Théorème. *Si une trajectoire* $\gamma : q = \varphi(t)$ *des équations de Lagrange* $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$ *s'écrit dans les coordonnées locales* Q, t $(Q = Q(q, t))$ *sous la forme* $\gamma : Q = \Phi(t)$, *alors la fonction* $\Phi(t)$ *vérifie les équations de Lagrange* $\frac{d}{dt}\frac{\partial L'}{\partial \dot{Q}} = \frac{\partial L'}{\partial Q}$, *où* $L'(Q, \dot{Q}, t) = L(q, \dot{q}, t)$:

Démonstration. La trajectoire γ est extrémale : $\delta \int_\gamma L(q, \dot{q}, t)\,dt = 0$. Donc $\delta \int_\gamma L'(Q, \dot{Q}, t)\,dt = 0$ et $\Phi(t)$ vérifie les équations de Lagrange, c.q.f.d.

B. Mouvements, rotations, mouvements de translation. Considérons le cas important où q est le rayon vecteur cartésien du point

mobile par rapport à un système de coordonnées inertial k (que nous appellerons *fixe*) et Q le rayon vecteur cartésien du même point par rapport à un système de coordonnées *mobile* K.

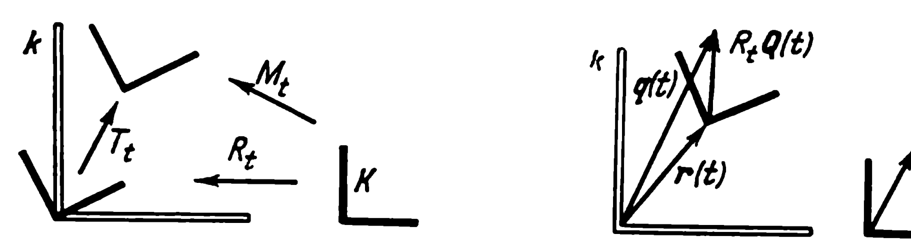


Fig. 103. Le mouvement M_t est le produit d'une rotation R_t par une translation T_t.

Fig. 104. Rayon vecteur d'un point par rapport aux systèmes de coordonnées immobile (q) et mobile (Q).

Définition. Soient k, K deux espaces vectoriels euclidiens orientés.

On appelle *mouvement de* K par rapport à k une application

$$M_t : K \to k,$$

différentiable par rapport à t préservant la métrique et l'orientation (fig. 103).

Définition. Un mouvement M_t est une *rotation* s'il transporte l'origine des coordonnées K dans l'origine des coordonnées k, i.e. si M_t est un opérateur linéaire.

Théorème. *Tout mouvement M_t se décompose de façon unique en un produit d'une rotation $R_t : K \to k$ par une translation $T_t : k \to k$:*

$$M_t = T_t R_t,$$

où $T_t q = q + r(t)$ $(q, r \in k)$.

Démonstration. Posons $r(t) = M_t O$, $R_t = T_t^{-1} M_t$. Alors $R_t O = O$, c.q.f.d.

Définition. Un mouvement M_t est un mouvement *de translation* si l'application $R_t : K \to k$ qui lui correspond ne dépend pas de t: $R_t = R_0 = R$, $M_t Q = RQ + r(t)$.

Nous dirons que k est *un système de coordonnées fixe*, K *un système de coordonnées mobile*, $q(t) \in k$ *le rayon vecteur du point mobile par rapport à* k, $Q(t)$ *le rayon vecteur du point mobile par rapport à* K, si (fig. 104)

$$q(t) = M_t Q(t) = R_t Q(t) + r(t). \tag{1}$$

Attention ! Ne pas confondre le vecteur $R_t Q(t) \in k$ et le vecteur $Q(t) \in K$: ils sont situés dans des espaces différents.

C. Composition des vitesses. Exprimons maintenant la « vitesse absolue » $\dot{q}$ en fonction du mouvement relatif de $Q(t)$ et du mouve-

ment M_t du système de coordonnées. La différentiation de (1) par rapport à t nous donne l'expression de la composition des vitesses

$$\dot{q} = \dot{R}Q + R\dot{Q} + \dot{r}. \qquad (2)$$

Interprétons les trois termes figurant dans (2) d'abord sur des cas particuliers.

Cas d'un mouvement de translation ($\dot{R} = 0$). L'équation (2) donne $\dot{q} = R\dot{Q} + \dot{r}$. Autrement dit est démontré le

Théorème. *Si le système mobile K se déplace en mouvement de translation par rapport à k, la vitesse absolue est égale à la somme de la vitesse relative et de la vitesse de K:*

$$v = v' + v_0, \qquad (3)$$

où

$v = \dot{q} \in k$ *est la vitesse absolue,*

$v' = R\dot{Q} \in k$ *la vitesse relative (qu'il ne faut pas confondre avec* $\dot{Q} \in K$!),

$v_0 = \dot{r} \in k$ *la vitesse du système de coordonnées mobile.*

D. Vitesse angulaire. Lorsque le système K est en rotation il n'est pas aussi simple d'établir une relation entre les vitesses relative et absolue. Considérons d'abord le cas où le point matériel est fixe par rapport à K (i.e. $\dot{Q} = 0$) et le système K est en mouvement de rotation (i.e. $r = 0$). Dans ce cas on dit que le mouvement du point $q(t)$ est une *rotation d'entraînement.*

E x e m p l e. *Rotation de vitesse angulaire constante* $\omega \in k$. Soit $U(t): k \to k$ une rotation de l'espace k d'angle $|\omega| t$ autour d'un

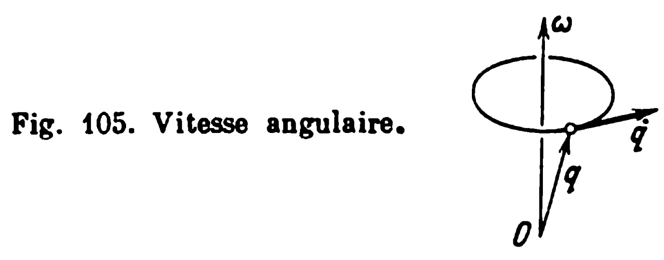


Fig. 105. Vitesse angulaire.

axe ω. Alors $R(t) = U(t) R(0)$ est une *rotation uniforme de K de vitesse angulaire* ω.

Il est évident que dans ce cas la vitesse d'entraînement du point q est donnée par la formule (fig. 105):

$$\dot{q} = [\omega, q].$$

Revenons maintenant au cas général de la rotation de K ($r = 0$, $\dot{Q} = 0$).

Théorème. *A chaque date t il existe un vecteur* $\omega(t) \in k$ *en fonction duquel la vitesse d'entraînement s'exprime au moyen de la formule*

$$\dot{q} = [\omega, q] \quad \forall q \in k. \tag{4}$$

Le vecteur ω est appelé *vitesse angulaire instantanée*; on voit qu'il est défini de façon unique par l'égalité (4).

Corollaire. *Soit un solide K en rotation autour d'un point fixe O de l'espace k. Alors à tout instant t il existe un axe instantané de rotation: une droite du solide passant par le point O telle que les vitesses de ses points soient nulles à l'instant t. Les vitesses des autres points sont perpendiculaires à cette droite et proportionnelles à la distance qui sépare ces points de cette droite.*

Dans l'espace k, l'axe instantané de rotation est défini par son vecteur ω; dans K le vecteur correspondant est désigné par $\Omega = R^{-1}\omega \in K$; Ω est par définition le *vecteur vitesse angulaire par rapport au solide.*

Exemple. La vitesse angulaire de la Terre est dirigée du centre vers le pôle Nord et égale à $\frac{2\pi}{3600\cdot 24}$ s$^{-1} \approx 7{,}3\cdot 10^{-5}$ s^{-1}.

Démonstration du théorème. En vertu de (2) on a

$$\dot{q} = \dot{R}Q.$$

En exprimant donc Q en fonction de q, on obtient $\dot{q} = \dot{R}R^{-1}q = Aq$, où $A = \dot{R}R^{-1}: k \to k$ est un opérateur linéaire de k dans k.

Lemme 1. *L'opérateur A est antisymétrique*: $A' + A = 0$.

Démonstration. Comme $R: K \to k$ est un opérateur orthogonal d'un espace euclidien dans un autre, son adjoint est confondu avec son inverse, $R' = R^{-1}: k \to K$. En dérivant l'expression $RR' = E$ par rapport à t on obtient

$$\dot{R}R' + R\dot{R}' = 0. \quad \dot{R}R^{-1} + (\dot{R}R^{-1})' = 0, \text{ c.q.f.d.}$$

Lemme 2. *Tout opérateur antisymétrique A de l'espace euclidien orienté* $\mathbf{R}^3$ *est opérateur d'un produit vectoriel par un vecteur fixe*:

$$Aq = [\omega, q] \quad \forall q \in \mathbf{R}^3.$$

Démonstration. Les opérateurs antisymétriques de $\mathbf{R}^3 \to \mathbf{R}^3$ forment un espace vectoriel de dimension 3, puisque les matrices antisymétriques 3×3 sont définies par leurs éléments subdiagonaux.

L'opérateur de produit vectoriel par ω est linéaire et antisymétrique. Les opérateurs de produit vectoriel par tous les vecteurs ω

de $\mathbf{R}^3$ forment un sous-espace vectoriel de l'espace des opérateurs antisymétriques.

Ce sous-espace est de dimension trois. Donc le sous-espace formé par les produits vectoriels est confondu avec l'espace des opérateurs antisymétriques, c.q.f.q.

Fin de la démonstration du théorème. En vertu des lemmes 1 et 2

$$\dot{q} = Aq = [\omega, q], \text{ c.q.f.d.}$$

En coordonnées cartésiennes, l'opérateur A est défini par une matrice antisymétrique; soient $\pm\omega_{1,2,3}$ ses éléments:

$$A = \begin{Vmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{Vmatrix}.$$

Dans cette notation, le vecteur $\omega = \omega_1 e_1 + \omega_2 e_2 + \omega_3 e_3$ sera un vecteur propre associé à la valeur propre 0. En appliquant A au vecteur $q = q_1 e_1 + q_2 e_2 + q_3 e_3$ on déduit immédiatement

$$Aq = [\omega, q].$$

E. Vitesse d'entraînement. *Cas d'une rotation pure.* Supposons maintenant que le système K est en rotation ($r = 0$) et que le point

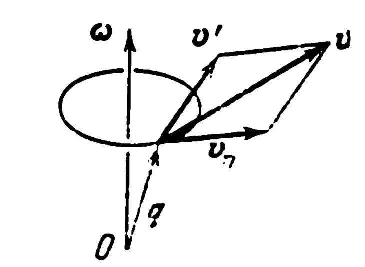


Fig. 106. Somme des vitesses.

matériel se déplace par rapport à K ($\dot{Q} \neq 0$). De (2) il vient (fig. 106)

$$\dot{q} = \dot{R}Q + R\dot{Q} = [\omega, q] + v'.$$

i.e. on a démontré le

Théorème. *Si le système K tourne autour de $O \in k$, la vitesse absolue est égale à la somme de la vitesse relative et de la vitesse d'entraînement de la rotation:*

$$v = v' + v_{\text{entr.}},$$

où

$$\left.\begin{array}{l} v = \dot{q} \in k \text{ est la vitesse absolue,} \\ v' = R\dot{Q} \in k \text{ la vitesse relative,} \\ v_{\text{entr.}} = \dot{R}Q = [\omega, q] \in k \text{ la vitesse d'entraînement de la rotation.} \end{array}\right\} \quad (5)$$

Voyons enfin un cas général qui peut être ramené aux deux précédents par introduction d'un système auxiliaire K_1 en mouvement de translation par rapport à k et par rapport auquel K est en rotation autour de $O \in K_1$. De la formule (2) on peut également déduire que

$$v = v' + v_{\text{entr.}} + v_0,$$

où

$v = \dot{q} \in k$ est la vitesse absolue,

$v' = \dot{R}Q \in k$ la vitesse relative,

$v_{\text{entr.}} = \dot{R}Q = [\omega, \; q - r] \in k$ la vitesse d'entraînement de la rotation,

$v_0 = \dot{r} \in k$ la vitesse du système de coordonnées mobile.

E x e r c i c e. Montrer que la vitesse angulaire d'un solide ne dépend pas du choix de l'origine d'un système de coordonnées mobile K par rapport à ce solide.

E x e r c i c e. Montrer que le mouvement le plus général d'un solide est le mouvement hélicoïdal, i.e. le produit d'une rotation d'angle φ autour d'un axe par une h-translation le long de cet axe.

E x e r c i c e. Une montre repose sur une table. Trouver la vitesse angulaire de la petite aiguille: a) par rapport à la Terre; b) par rapport à un système de coordonnées inertial.

I n d i c a t i o n. Si l'on a trois systèmes de coordonnées k, K_1 et K_2, la vitesse angulaire de K_2 par rapport à k est égale à la somme des vitesses angulaires de K_1 par rapport à k et de K_2 par rapport à K_1. En effet

$$(E + A_1t + \ldots)(E + A_2t + \ldots) = E + (A_1 + A_2)t + \ldots$$

§ 27. Forces d'inertie. Force de Coriolis

Les équations de mouvement dans un système de coordonnées non inertial se distinguent des équations de mouvement dans un système inertial par la présence de termes supplémentaires appelés forces d'inertie. Ceci permet de déterminer expérimentalement si un mouvement est inertial ou non (exemple: la rotation de la Terre autour de son axe).

A. Système de coordonnées en translation.

Théorème. *Dans un système de coordonnées K en mouvement de translation par rapport à un système inertial k, les systèmes mécaniques se déplacent comme si le système de coordonnées était inertial, et sur chaque point de masse m agissait une « force d'inertie » supplémentaire $F = -m\ddot{r}$, où $\ddot{r}$ est l'accélération du système K.*

D é m o n s t r a t i o n. Si $Q = q - r(t)$, alors $m\ddot{Q} = m\ddot{q} - m\ddot{r}$. L'influence du mouvement de translation du système de coordonnées se traduit donc par l'apparition d'une force supplémentaire d'un champ de forces homogène $-mW$, où W est l'accélération de l'origine des coordonnées, c.q.f.d.

Exemple 1. Au départ, la fusée de la figure 107 est douée d'une accélération ascendante $\ddot{r}$. Donc le système de coordonnées K lié à la fusée n'est pas inertial et son passager peut mettre en évidence le champ $-mW$ et mesurer la force d'inertie avec une balance à ressort par exemple. Dans ce cas la force d'inertie porte le nom de *surcharge*.

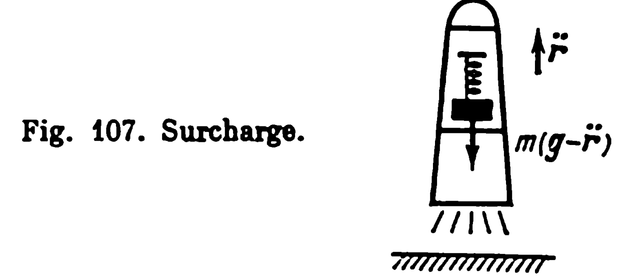


Fig. 107. Surcharge.

Exemple 2. Dans les plongeons de haut vol les plongeurs ont une accélération descendante g. Donc la somme de la force d'inertie et du poids est nulle. La balance à ressort indiquant un poids nul, cet état est appelé *état d'apesanteur*. De la même façon on observe l'état d'apesanteur en vol balistique libre d'un satellite, car la force d'inertie est opposée à la pesanteur de la Terre.

Exemple 3. Si le point d'attache d'un pendule possède une accélération $W(t)$, le pendule se déplace comme si l'accélération de la force de pesanteur g était variable et égale à $g - W(t)$.

B. Système de coordonnées en rotation. Soit $R_t : K \to k$ la rotation d'un système de coordonnées K par rapport à un système de coordonnées fixe k. Désignons par $Q(t) \in K$ le rayon vecteur du point en mouvement par rapport au système de coordonnées mobile et par $q(t) = R_t Q(t) \in k$ le rayon vecteur de ce point par rapport au système fixe. Comme dans le § 26, désignons par Ω le vecteur vitesse angulaire de rotation dans le système mobile.

Supposons que par rapport au système k le mouvement du point q obéisse à l'équation de Newton $m\ddot{q} = f(q, \dot{q})$.

Théorème. *Par rapport au système de coordonnées en rotation le point Q de masse m se meut comme s'il était soumis à l'action de trois « forces d'inertie » supplémentaires:*

la force d'inertie de rotation: $m[\dot{\Omega}, Q]$,

la force de Coriolis: $2m[\Omega, \dot{Q}]$,

la force centrifuge: $m[\Omega, [\Omega, Q]]$.

Donc

$$m\ddot{Q} = F - m[\dot{\Omega}, Q] - 2m[\Omega, \dot{Q}] - m[\Omega, [\Omega, Q]],$$

où

$$RF(Q, \dot{Q}) = f(RQ, (R\dot{Q})).$$

La première de ces forces d'inertie s'observe seulement dans le cas de rotations non uniformes; les deux autres interviennent dans les rotations uniformes.

La force centrifuge (fig. 108) est toujours dirigée de l'axe instantané de rotation Ω vers l'extérieur et est égale en module à $|\Omega|^2 r$, où r est la distance à cet axe. Cette force ne dépend pas de la vitesse

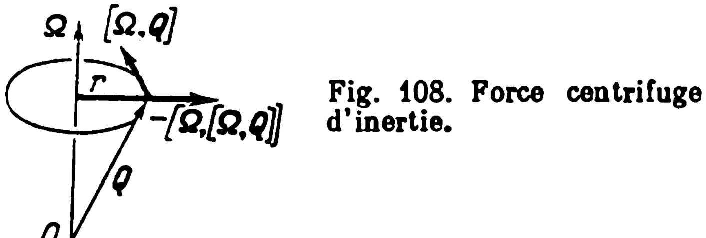


Fig. 108. Force centrifuge d'inertie.

du mouvement relatif et agit même sur les corps en repos par rapport au système de coordonnées K.

La force de Coriolis dépend de la vitesse $\dot{Q}$. Dans l'hémisphère boréal elle dévie à droite tout mouvement s'effectuant sur Terre et vers l'est tout corps en chute libre.

D é m o n s t r a t i o n d u t h é o r è m e. On remarquera que pour tout vecteur $X \in K$ on a $\dot{R}X = R\,[\Omega, X]$. En effet, en vertu du § 26 $\dot{R}X = [\omega, x] - [R\Omega, RX]$. Ceci est égal à $R\,[\Omega, X]$, puisque l'opérateur R préserve la métrique et l'orientation et par conséquent le produit vectoriel.

De $q = RQ$ on déduit que $\dot{q} = \dot{R}Q + R\dot{Q} = R\,(\dot{Q} + [\Omega, Q])$. En différentiant une seconde fois on obtient

$$\ddot{q} = \dot{R}\,(\dot{Q} + [\Omega, Q]) + R\,(\ddot{Q} + [\dot{\Omega}, Q] + [\Omega, \dot{Q}]) =$$
$$= R\,([\Omega, (\dot{Q} + [\Omega, Q])] + \ddot{Q} + [\dot{\Omega}, Q] + [\Omega, \dot{Q}]) =$$
$$= R\,(\ddot{Q} + 2\,[\Omega, \dot{Q}] + [\Omega, [\Omega, Q]] + [\dot{\Omega}, Q]), \text{ c.q.f.d.}$$

(Nous nous sommes une fois de plus servis de l'expression $\dot{R}X = R\,[\Omega, X]$, ici $X = \dot{Q} + [\Omega, Q]$.)

Etudions plus en détail l'influence de la rotation de la Terre sur les expériences en laboratoire. Comme la Terre tourne pratiquement uniformément, on peut supposer que $\dot{\Omega} = 0$. La force centrifuge a la plus grande valeur sur l'équateur où elle atteint $\frac{\Omega^2\rho}{g} \approx \frac{(7{,}3\cdot 10^{-5})^2\cdot 6{,}4\cdot 10^6}{9{,}8} \approx \frac{3}{1000}$ du poids. Cependant dans les limites du laboratoire elle varie peu et pour l'observer il faudrait voyager.

Donc en laboratoire, la rotation de la Terre ne se manifeste que par la force de Coriolis : par rapport à un système de coordonnées Q rattaché à la Terre on a avec une grande précision

$$\frac{d}{dt} m\dot{Q} = m g + 2m\,[\dot{Q}, \Omega]$$

(la force centrifuge est prise en considération dans g).

E x e m p l e 1. Une pierre est lâchée (sans vitesse initiale) dans un puits d'une profondeur de 250 m à la latitude de Léningrad ($\lambda = 60°$). Déterminer son écart par rapport à la verticale.

Résolvons l'équation

$$\ddot{Q} = g + 2[\dot{Q}, \Omega]$$

par approximations successives sachant que $\Omega \ll 1$. Posons (fig. 109)

$$Q = Q_1 + Q_2, \quad \text{où} \quad \dot{Q}_2(0) = Q_2(0) = 0 \quad \text{et} \quad Q_1 = Q_1(0) + \frac{gt^2}{2}.$$

Relativement à Q_2 nous obtenons

$$\ddot{Q}_2 = 2[gt, \Omega] + O(\Omega^2), \quad Q_2 \approx \frac{t^3}{3}[g, \Omega] \approx \frac{2t}{3}[h, \Omega], \quad h = \frac{gt^2}{2}.$$

On voit que la pierre s'écarte vers l'est environ de

$$\frac{2t}{3}|h||\Omega|\cos\lambda \approx \frac{2\cdot 7}{3}\cdot 250\cdot 7\cdot 10^{-5}\cdot\frac{1}{2}\text{ m} \approx 4\text{ cm}.$$

E x e r c i c e. Un obus tiré verticalement à Léningrad atteint une hauteur de 1 km. De combien sera-t-il dévié de l'axe du fût par la force de Coriolis?

E x e m p l e 2. *Le pendule de Foucault* (fig. 110).

Considérons les petites oscillations du pendule mathématique en tenant compte de la force de Coriolis. Soit e_x, e_y, e_z un système de coordonnées lié à

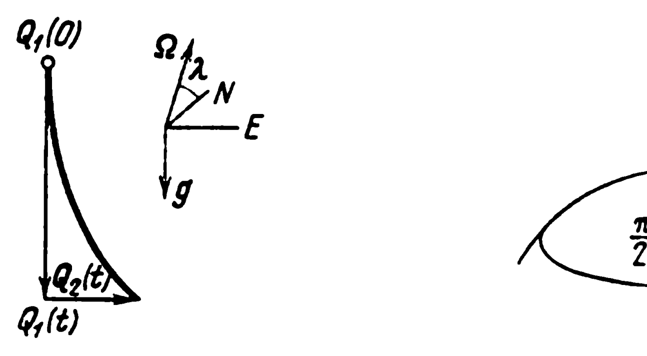


Fig. 109. Ecart dû à l'action de la force de Coriolis.

Fig. 110. Système de coordonnées pour l'étude du mouvement du pendule de Foucault.

la Terre: e_z est dirigé verticalement vers le haut, e_x et e_y sont situés dans un plan horizontal. Dans l'approximation des petites oscillations $\dot{z} = 0$ (en comparaison avec $\dot{x}$, $\dot{y}$), donc la composante horizontale de la force de Coriolis sera $2m\dot{y}\Omega_z e_x - 2m\dot{x}\Omega_z e_y$. D'où l'on déduit les équations de mouvement

$$\begin{cases} \ddot{x} = -\omega^2 x + 2\dot{y}\Omega_z, \\ \ddot{y} = -\omega^2 y - 2\dot{x}\Omega_z \end{cases} \quad (\Omega_z = |\Omega|\sin\lambda_0, \text{ où } \lambda_0 \text{ est la latitude}).$$

Si l'on pose $x + iy = w$, on obtient $\dot{w} = \dot{x} + i\dot{y}$, $\ddot{w} = \ddot{x} + i\ddot{y}$ et les deux équations se ramènent à une seule équation complexe

$$\ddot{w} + i2\Omega_z\dot{w} + \omega^2 w = 0.$$

Résolvons-la : $w = e^{\lambda t}$, $\lambda^2 + 2i\Omega_z\lambda + \omega^2 = 0$, $\lambda = -i\Omega_z \pm i\sqrt{\Omega_z^2+\omega^2}$. Or $\Omega_z^2 \ll \omega^2$. Donc $\sqrt{\Omega_z^2 + \omega^2} = \omega + O(\Omega_z^2)$, d'où, en négligeant Ω_z^2,

$$\lambda \approx -i\Omega_z \pm i\omega$$

ou avec la même précision

$$w = e^{-i\Omega_z t}(c_1 e^{i\omega t} + c_2 e^{-i\omega t}).$$

Pour $\Omega_z = 0$ on obtient les oscillations harmoniques habituelles du pendule sphérique. Nous constatons que l'influence de la force de Coriolis se traduit par une rotation de la figure avec une vitesse angulaire $-\Omega_z$, où $|\Omega_z| = |\mathbf{\Omega}| \sin \lambda_0$.

En particulier, si les conditions initiales correspondent à un mouvement plan ($y(0) = \dot{y}(0) = 0$), alors le plan des oscillations tournera à une vitesse angulaire $-\Omega_z$ par rapport au système de coordonnées lié à la Terre (fig. 111).

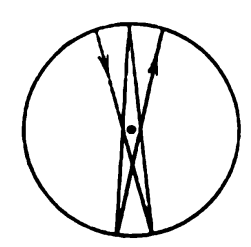


Fig. 111. Trajectoire du pendule de Foucault.

Au pôle, le plan des oscillations effectue un tour en 24 heures (ce plan est fixe par rapport à tout système de coordonnées ne tournant pas avec la Terre). A la latitude de Moscou (56°) le plan fera 0,83 tour en 24 heures, i.e. il pivotera de 12,5° à l'heure.

Exercice. La vitesse d'un fleuve est de 3 km/h. Pour quelle valeur du rayon de courbure d'un coude la force de Coriolis due à la rotation de la Terre est-elle supérieure à la force centrifuge définie par ce coude?

Réponse. Le rayon de courbure ne doit pas être inférieur à 10 km pour les fleuves des latitudes moyennes.

La solution de cet exercice explique d'ailleurs pourquoi les fleuves de l'hémisphère boréal (le cours moyen de la Volga par exemple) affouillent principalement leurs rives droites, alors qu'une rivière comme la Moskova avec ses méandres de faible rayon de courbure affouille tantôt la rive droite, tantôt la rive gauche.

§ 28. Le solide

Dans ce paragraphe on définit le solide, son tenseur d'inertie, son ellipsoïde d'inertie, ses moments et axes d'inertie.

A. Variété de configuration du solide.

Définition. Un *solide* est par définition un système de points matériels soumis à la liaison holonome

$$|x_i - x_j| = r_{ij} = \text{const}, \tag{1}$$

qui traduit le fait que la distance des points est constante.

Théorème. *La variété de configuration du solide est une variété de dimension six, plus exactement, le produit cartésien* $\mathbf{R}^3 \times SO(3)$ *de l'espace* $\mathbf{R}^3$ *par son groupe des rotations* $SO(3)$, *si seulement ce solide est composé de trois points non alignés.*

D é m o n s t r a t i o n. Soient x_1, x_2 et x_3 trois points du solide non alignés. Considérons un repère orthonormé direct dont le premier vecteur est orienté dans le sens de x_2-x_1, le second du côté de x_3 dans le plan x_1, x_2, x_3 (fig. 112). Des conditions $|x_i-x_j| = r_{ij}$ $(i = 1, 2, 3)$ il résulte que la position de tous les points du solide

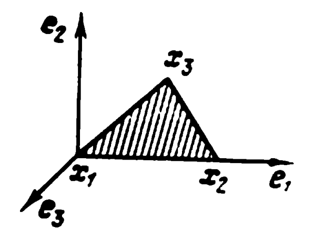


Fig. 112. Variété de configuration du solide.

est définie de façon unique par la position de x_1, x_2, x_3, cette dernière étant définie à son tour par la position du repère. Enfin l'espace des repères de $\mathbf{R}^3$ est $\mathbf{R}^3 \times SO(3)$, puisque chaque repère s'obtient à partir d'un repère fixe par une rotation et une translation *).

E x e r c i c e. Trouver l'espace de configuration d'un solide dont tous les points sont alignés.

Réponse: $\mathbf{R}^3 \times S^2$.

D é f i n i t i o n. Un *solide ayant un point fixe O* est un système de points matériels comportant outre (1) la liaison $x_1 = 0$.

De toute évidence sa variété de configuration est le groupe des rotations $SO(3)$.

B. Lois de conservation. Considérons le déplacement par inertie d'un solide libre en dehors d'un champ de forces. Un exemple (approximatif) est la culbute d'un engin spatial.

Ce système admet tous les mouvements de translation: ils préservent le lagrangien. En vertu du théorème de Noether existent trois premières intégrales: les trois composantes du vecteur impulsion. Autrement dit on a démontré le

Théorème. *Lorsqu'un solide est en mouvement libre, son centre d'inertie est en mouvement rectiligne uniforme.*

Ceci étant, on peut considérer un système de coordonnées inertial par rapport auquel le centre d'inertie est fixe. On obtient le

Corollaire. *Un solide libre tourne autour de son centre d'inertie comme si ce centre était assimilé à un point fixe O.*

Le problème s'est donc ramené à une rotation à trois degrés de liberté autour d'un point fixe O. Etudions en détail ce problème (sans supposer forcément que O est le centre d'inertie du solide).

*) En toute rigueur, l'espace de configuration d'un solide est $\mathbf{R}^3 \times O(3)$; quant à $\mathbf{R}^3 \times SO(3)$, ce n'est que l'une des deux composantes connexes de cette variété correspondant à une orientation définie du solide.

Le lagrangien est invariant par toutes les rotations autour de O. En vertu du théorème de Noether, il existe trois intégrales premières : les trois composantes du vecteur moment cinétique. L'énergie totale (elle se réduit ici à l'énergie cinétique : $E = T$) se conserve également. On a donc démontré le

Théorème. *Quand un solide est en mouvement autour d'un point fixe O, en l'absence de forces extérieures on dispose de quatre intégrales premières :* M_x, M_y, M_z, E.

Ce théorème permet sans effectuer le moindre calcul de tirer les conclusions nécessaires sur le mouvement.

La position et la vitesse du solide sont définies par un point de la variété de dimension six TSO (3), fibré tangent à la variété de configuration SO (3). Les intégrales premières M_x, M_y, M_z, E sont des fonctions définies sur TSO (3). On vérifie que dans le cas général (si le solide ne possède pas de symétrie spéciale) ces quatre fonctions sont indépendantes. Donc les quatre équations

$$M_x = C_1, \quad M_y = C_2, \quad M_z = C_3, \quad E = C_4 > 0$$

définissent une sous-variété V_c de dimension deux de TSO(3).

Cette variété est invariante : si les conditions initiales du mouvement définissent un point de V_c, alors le point de TSO(3) qui figure la position et la vitesse du solide restera sur V_c pendant toute la durée du mouvement.

C'est pourquoi la variété V_c admet un champ de vecteurs tangent (plus exactement le champ de vitesses du mouvement sur TSO (3)) ; pour $C_4 > 0$ ce champ ne peut posséder de points singuliers. On vérifie aisément que V_c est compacte (utiliser E) et orientable (puisque TSO (3) l'est) *).

En topologie on montre que toutes les variétés compactes orientables et connexes de dimension deux sont des sphères à n anses, $n \geqslant 0$ (fig. 113). Parmi elles seul le tore ($n = 1$) admet un champ de vecteurs tangent n'ayant pas de points singuliers.

Donc la variété invariante V_c est un tore de dimension deux (ou plusieurs tores).

---

*) On démontre aisément les propositions suivantes.

1. Soient $f_1, \ldots, f_k : M \to \mathbf{R}$ des fonctions sur une variété orientée M. Considérons l'ensemble V défini par les équations $f_1 = c_1, \ldots, f_k = c_k$. Supposons que les gradients de $f_1, \ldots, f_k$ sont linéairement indépendants en chaque point de V. Alors V est orientable.
2. Le produit cartésien de variétés orientables est orientable.
3. Le fibré tangent TSO(3) est le produit cartésien $\mathbf{R}^3 \times$ SO(3).

On dit qu'une variété est *parallélisable* si son fibré tangent est un produit cartésien. Le groupe SO(3) (comme d'ailleurs tout groupe de Lie) est parallélisable.

4. Une variété parallélisable est orientable.

De 1 à 4 découle l'orientabilité de SO(3), TSO(3) et de V_c.

Nous verrons dans la suite que sur ce tore on peut choisir des coordonnées angulaires φ_1, φ_2 (mod 2π) telles que le mouvement du point représentatif sur V_c soit défini par les équations $\dot{\varphi}_1 = \omega_1(c)$, $\dot{\varphi}_2 = \omega_2(c)$.

En d'autres termes, la rotation du solide est la composition de deux mouvements périodiques, en général de périodes différentes:

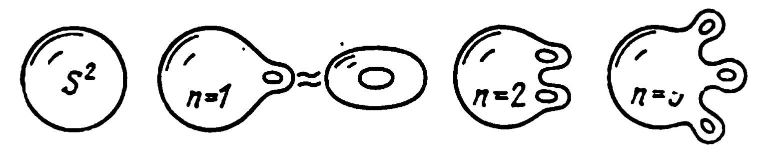
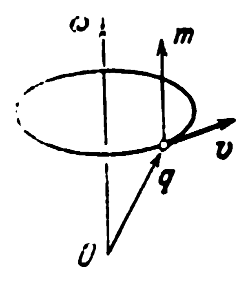


Fig. 113. Variétés orientées, connexes, compactes, de dimension deux.

si les fréquences ω_1 et ω_2 sont incommensurables, alors le solide ne revient jamais à un état passé du mouvement. La valeur des fréquences ω_1 et ω_2 dépend des conditions initiales c.

C. Opérateur d'inertie. Passons maintenant à l'aspect quantitatif de la théorie et introduisons les *notations* suivantes. Soit k un système de coordonnées fixe et K un système de coordonnées tournant

Fig. 114. Rayon vecteur, vecteur vitesse, vitesse angulaire et moment cinétique d'un point du solide par rapport à l'espace.



avec le solide autour d'un point O; par rapport à ce système le solide est au repos. Tout vecteur de l'espace K est susceptible d'être transporté dans l'espace k par un opérateur R. Les vecteurs qui se correspondent seront désignés par une même lettre: cette lettre sera majuscule si le vecteur appartient à K et minuscule s'il appartient à k. Ainsi (fig. 114):

$\boldsymbol{q} \in k$ est le rayon vecteur du point en mouvement par rapport à l'espace,

$\boldsymbol{Q} \in K$, le rayon vecteur par rapport au solide, $\boldsymbol{q} = R\boldsymbol{Q}$,

$\boldsymbol{v} = \dot{\boldsymbol{q}} \in k$, le vecteur vitesse du point par rapport à l'espace,

$\boldsymbol{V} \in K$, le même vecteur par rapport au solide, $\boldsymbol{v} = R\boldsymbol{V}$,

$\boldsymbol{\omega} \in k$, la vitesse angulaire du point par rapport à l'espace,

$\boldsymbol{\Omega} \in K$, la vitesse angulaire par rapport au solide, $\boldsymbol{\omega} = R\boldsymbol{\Omega}$,

$\boldsymbol{m} \in k$, le moment cinétique par rapport à l'espace,

$\boldsymbol{M} \in K$, le moment cinétique par rapport au solide, $\boldsymbol{m} = R\boldsymbol{M}$.

Comme l'opérateur $R: K \to k$ conserve la métrique et l orientation, il conserve également les produits scalaires et vectoriels.

Par définition de la vitesse angulaire (§ 26)

$$v = [\omega, q].$$

Par définition du moment cinétique d'un point de masse m par rapport à O

$$\boldsymbol{m} = [q, mv] = m\,[q, [\omega, q]].$$

Donc

$$\boldsymbol{M} = m\,[Q, [\Omega, Q]].$$

Et l'on obtient un opérateur linéaire A transformant Ω en $\boldsymbol{M}$:

$$A : K \to K, \quad A\Omega = \boldsymbol{M}.$$

Cet opérateur dépend encore du point (Q) du solide et de sa masse (m).

Lemme. *L'opérateur A est symétrique.*

D é m o n s t r a t i o n. Quels que soient $X \in K$, $Y \in K$, en vertu de la relation $([a, b], c) = ([c, a], b)$ on a

$$(AX, Y) = m\,(Q, [X, Q]], Y) = m\,([Y, Q], [X, Q]),$$

or cette dernière expression est symétrique en X et Y, c.q.f.d.

En remplaçant X et Y par le vecteur vitesse angulaire Ω et en remarquant que $[\Omega, Q]^2 = V^2 = v^2$, on obtient le

Corollaire. *L'énergie cinétique d'un point du solide est une forme quadratique en le vecteur vitesse angulaire Ω, plus exactement*

$$T = \frac{1}{2}(A\Omega, \Omega) = \frac{1}{2}(\boldsymbol{M}, \Omega).$$

L'opérateur symétrique A porte le nom d'*opérateur* (ou tenseur) *d'inertie du point Q*.

Si le solide est constitué de points Q_i de masses m_i, alors en sommant on obtient le

Théorème. *Le moment cinétique $\boldsymbol{M}$ du solide par rapport à un point fixe O dépend linéairement de la vitesse angulaire Ω, i.e. il existe un opérateur linéaire $A : K \to K$, $A\Omega = \boldsymbol{M}$. L'opérateur A est symétrique.*

L'énergie cinétique du solide est une forme quadratique en la vitesse angulaire Ω,

$$T = \frac{1}{2}(A\Omega, \Omega) = \frac{1}{2}(\boldsymbol{M}, \Omega).$$

D é m o n s t r a t i o n. Par définition, le moment cinétique du solide est égal à la somme des moments de ses points :

$$\boldsymbol{M} = \sum_i \boldsymbol{M}_i = \sum_i A_i\Omega = A\Omega, \text{ où } A = \sum_i A_i.$$

Comme l'opérateur d'inertie A_i de chaque point est symétrique en vertu du lemme, l'opérateur A l'est également. S'agissant de

l'énergie cinétique on a par définition

$$T=\sum_i T_i=\sum_i \frac{1}{2}(\boldsymbol{M}_i, \Omega)=\frac{1}{2}(\boldsymbol{M}, \Omega)=\frac{1}{2}(A\Omega, \Omega), \text{ c.q.f.d.}$$

D. Axes d'inertie. Comme tout opérateur symétrique, l'opérateur A possède trois directions propres deux à deux orthogonales. Soient $e_1, e_2, e_3 \in K$ leurs vecteurs unitaires, I_1, I_2, I_3 les valeurs propres. Dans la base e_i l'opérateur d'inertie et l'énergie cinétique s'écrivent sous une forme particulièrement simple:

$$A_i=I_i\Omega_i,$$
$$T=\frac{1}{2}(I_1\Omega_1^2+I_2\Omega_2^2+I_3\Omega_3^2).$$

Les axes e_i s'appellent *axes d'inertie* du solide au point O.

De toute évidence, si les valeurs I_1, I_2, I_3 ne sont pas toutes distinctes, alors les axes d'inertie e_i ne sont pas définis de façon unique. Interprétons en détail les valeurs propres I_1, I_2, I_3.

Théorème. *Lorsqu'un solide ayant un point fixe O tourne autour d'un axe e à la vitesse angulaire $\Omega = \Omega e$ $(\Omega = |\Omega|)$, son énergie cinétique est égale à*

$$T=\frac{1}{2}I_e\Omega^2, \quad \text{où} \quad I_e=\sum_i m_i r_i^2$$

et r_i désigne la distance du i-ème point à l'axe e (fig. 115).

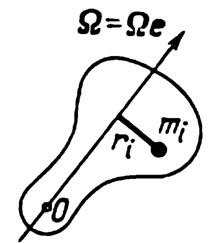


Fig. 115. Energie cinétique du solide en rotation autour d'un axe.

D é m o n s t r a t i o n. Par définition $T=\frac{1}{2}\sum_i m_i v_i^2$, or $|v_i|= \Omega r_i$, donc $T=\frac{1}{2}\left(\sum_i m_i r_i^2\right)\Omega^2$, c.q.f.d.

La quantité I_e dépend de l'orientation e de l'axe de rotation Ω par rapport au solide.

D é f i n i t i o n. I_e est appelé *moment d'inertie du solide par rapport à l'axe e*:

$$I_e=\sum_i m_i r_i^2.$$

En comparant les deux expressions de T on obtient le

Corollaire. *Les valeurs propres I_i de l'opérateur d'inertie A sont les moments d'inertie du solide par rapport aux axes d'inertie e_i.*

E. Ellipsoïde d'inertie. Voyons comment le moment d'inertie I_e dépend de l'orientation de l'axe e par rapport au solide ; pour cela considérons les vecteurs $e/\sqrt{I_e}$, où le vecteur unitaire e parcourt la sphère unité.

Théorème. *Les vecteurs $e/\sqrt{I_e}$ forment un ellipsoïde dans K.*

Démonstration. Si $\Omega = e/\sqrt{I_e}$, alors la forme quadratique $T = \frac{1}{2}(A\Omega, \Omega)$ est égale à $\frac{1}{2}$. Donc $\{\Omega\}$ est un ensemble de niveau d'énergie de la forme quadratique définie positive, i.e. un

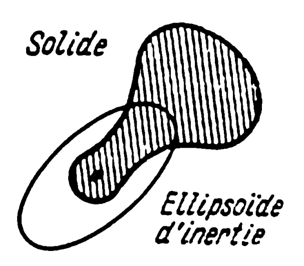


Fig. 116. Ellipsoïde d'inertie.

ellipsoïde, c.q.f.d. On peut dire que cet ellipsoïde est composé des vecteurs vitesse angulaire Ω pour lesquels l'énergie cinétique est égale à $\frac{1}{2}$.

Définition. L'ellipsoïde $\{\Omega : (A\Omega, \Omega) = 1\}$ est appelé *ellipsoïde d'inertie du solide* au point O (fig. 116).

Dans les axes d'inertie e_i son équation est de la forme

$$I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2 = 1.$$

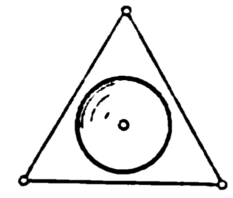
Donc les *axes principaux de l'ellipsoïde d'inertie sont dirigés suivant les axes d'inertie et leurs longueurs sont inversement proportionnelles à $\sqrt{I_i}$.*

Remarque. Si le solide est allongé dans le sens d'un axe quelconque, le moment d'inertie par rapport à cet axe est peu élevé, donc l'ellipsoïde d'inertie est également allongé dans le sens de cet axe : ainsi l'ellipsoïde d'inertie imite certains traits du solide.

Si le solide possède un axe de symétrie d'ordre k passant par O (i.e. il se confond avec lui-même par une rotation de $2\pi/k$ autour de cet axe), alors son ellipsoïde d'inertie possède la même symétrie par rapport à cet axe. Or les ellipsoïdes à trois axes ne possèdent pas d'axe de symétrie d'ordre $k > 2$. Donc tout axe de symétrie d'ordre $k > 2$ est un axe de rotation de l'ellipsoïde d'inertie et par conséquent son axe principal.

E x e m p l e. L'ellipsoïde d'inertie de trois points de masse m, disposés aux sommets d'un triangle équilatéral de centre O est un ellipsoïde de révolution autour de la normale en O au plan du triangle (fig. 117).

Fig. 117. Ellipsoïde d'inertie du triangle équilatéral.



S'il existe plusieurs tels axes, alors l'ellipsoïde d'inertie est une boule et tout axe est principal.

E x e r c i c e. Mener par le centre d'un cube une droite telle que la somme des carrés des distances de cette droite aux sommets du cube soit : a) maximale ; b) minimale.

Remarquons maintenant que l'ellipsoïde d'inertie (ou l'opérateur d'inertie ou les moments d'inertie I_1, I_2, I_3) définissent entièrement les propriétés giratoires de notre solide : si nous considérons deux solides ayant les mêmes ellipsoïdes d'inertie, alors pour des conditions initiales identiques ces solides auront le même mouvement (puisqu'ils auront même lagrangien $L = T$).

Donc du point de vue de la dynamique de rotation autour de O, l'*espace de tous les solides est de dimension trois* quel que soit le nombre de points dont ils sont composés.

Nous pouvons même considérer un « solide continu de densité $\rho(Q)$ », i.e. la limite pour $\Delta Q \to 0$ d'une suite de solides constitués

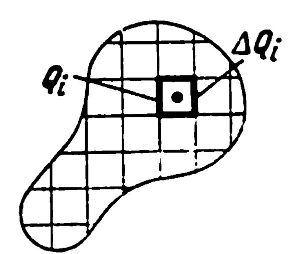


Fig. 118. Solide plein.

d'un nombre fini de points Q_i de masses $\rho(Q_i)\,\Delta Q_i$ (fig. 118) ou, ce qui revient au même, un solide quelconque de moments d'inertie

$$I_e = \iiint \rho(Q)\, r^2(Q)\, dQ,$$

où r est la distance de Q à l'axe e.

E x e m p l e. Trouver les axes et les moments d'inertie d'une lame homogène $|x| \leqslant a$, $|y| \leqslant b$, $z = 0$ par rapport à O.

S o l u t i o n. Comme la lame possède trois plans de symétrie, l'ellipsoïde d'inertie admet les mêmes plans de symétrie et donc les axes de symétrie x, y, z. Par ailleurs

$$I_x = \int_{-a}^{a}\int_{-b}^{b} x^2 \rho\, dx\, dy = \frac{ma^2}{3},$$

de la même manière

$$I_y = \frac{mb^2}{3};$$

et de toute évidence $I_z = I_x + I_y$.

Exercice. Démontrer que les moments d'inertie de tout solide vérifient les inégalités du triangle

$$I_3 \leqslant I_2 + I_1, \quad I_2 \leqslant I_1 + I_3, \quad I_1 \leqslant I_2 + I_3,$$

l'égalité n'étant réalisée que dans le cas d'un solide plan.

Exercice. Trouver les axes et les moments d'inertie d'un ellipsoïde homogène de masse m et de demi-axes a, b, c par rapport à son centre O.
Indication. Traiter d'abord le cas d'une boule.

Exercice. Démontrer le théorème de Steiner:
Les moments d'inertie d'un solide quelconque par rapport à deux axes parallèles dont l'un passe par le centre d'inertie sont liés par la relation

$$I = I_0 + mr^2,$$

où m est la masse du solide, r la distance entre les axes, I_0 le moment d'inertie par rapport à l'axe qui passe par le centre d'inertie.

Donc le moment d'inertie par rapport à un axe passant par le centre d'inertie est inférieur au moment d'inertie par rapport à un axe parallèle quelconque.

Exercice. Trouver les axes et les moments d'inertie d'un tétraèdre homogène par rapport à son sommet.

Exercice. Tracer le vecteur du moment cinétique $\boldsymbol{M}$ d'un solide dont l'ellipsoïde d'inertie est donné et tournant à une vitesse angulaire donnée $\boldsymbol{\Omega}$.

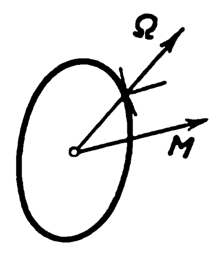


Fig. 119. Vitesse angulaire, ellipsoïde d'inertie et moment cinétique.

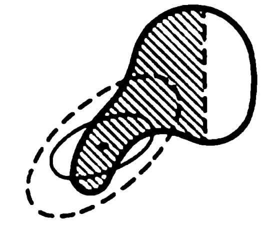


Fig. 120. Comportement des moments d'inertie lorsque le solide diminue de volume.

Réponse. $\boldsymbol{M}$ est de même sens que la normale à l'ellipsoïde d'inertie en un point d'intersection avec l'ellipsoïde de l'axe $\boldsymbol{\Omega}$ (fig. 119).

Exercice. Comment varient les principaux moments d'inertie d'un solide ayant un point fixe O dont on a découpé un morceau (fig. 120)?
Réponse. Les trois moments diminuent.
Indication. Cf. § 24.

Exercice. A un solide de moments d'inertie $I_1 > I_2 > I_3$ on ajoute une masse ε au point $Q = x_1 e_1 + x_2 e_2 + x_3 e_3$. Trouver la variation de I_1 et e_1 à $O(\varepsilon^2)$ près.

S o l u t i o n. Le centre d'inertie se déplace d'une distance de l'ordre de ε. Donc les moments d'inertie du solide initial par rapport aux axes parallèles passant par le nouveau et l'ancien centre d'inertie diffèrent d'une quantité de l'ordre de ε^2. D'autre part, l'adjonction de la masse modifie le moment d'inertie par rapport à tout axe fixe d'une quantité de l'ordre de ε. Si donc dans les calculs on admet une erreur de l'ordre de $O(\varepsilon^2)$, on peut négliger le déplacement du centre d'inertie.

Donc après adjonction de la petite masse ε l'énergie cinétique prend la forme

$$T = T_0 + \frac{1}{2}\varepsilon\,[\Omega, Q]^2 + O(\varepsilon^2)$$

où $T_0 = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)$ est l'énergie cinétique du solide initial. Cherchons la valeur propre $I_1(\varepsilon)$ et le vecteur propre $e_1(\varepsilon)$ de l'opérateur d'inertie sous forme d'une série de Taylor en ε. En égalant les coefficients en ε dans la relation $A(\varepsilon)\,e_1(\varepsilon) = I_1(\varepsilon)\,e_1(\varepsilon)$ il vient à $O(\varepsilon^2)$ près:

$$I_1(\varepsilon) \approx I_1 + \varepsilon(x_2^2 + x_3^2), \qquad e_1(\varepsilon) \approx e_1 + \varepsilon\left(\frac{x_1x_2}{I_2 - I_1}e_2 + \frac{x_1x_3}{I_3 - I_1}e_3\right).$$

De l'expression de $I_1(\varepsilon)$ on voit que les principaux moments d'inertie varient (en première approximation sur ε) comme si le centre et les axes d'inertie étaient fixes. L'expression de $e_1(\varepsilon)$ montre comment varie le sens des axes principaux: le plus proche demi grand axe de l'ellipsoïde d'inertie se rapproche du point ajouté, tandis que le demi petit axe s'en éloigne. Par ailleurs, l'adjonction d'une petite masse à l'un des principaux plans de l'ellipsoïde d'inertie fait pivoter les deux axes contenus dans ce plan et ne modifie pas le sens du troisième axe. Les différences de moments d'inertie dans le dénominateur sont dues au fait que les axes principaux d'un ellipsoïde de révolution sont indéterminés. Si l'ellipsoïde d'inertie est proche d'un ellipsoïde de révolution (par exemple $I_1 \approx I_2$), l'adjonction d'une petite masse peut faire fortement tourner les axes principaux e_1 et e_2 dans le plan qu'ils engendrent.

§ 29. Equations d'Euler. Equations du mouvement de Poinsot

On étudie ici le mouvement du solide autour d'un point fixe en l'absence de forces extérieures et partant le mouvement d'un solide libre. Nous verrons que ce mouvement possède deux fréquences.

A. Equation d'Euler. Considérons le mouvement d'un solide autour d'un point fixe O. Soient $\boldsymbol{M}$ le vecteur du moment cinétique du solide par rapport à O et par rapport au solide, $\boldsymbol{\Omega}$ le vecteur vitesse angulaire par rapport au solide, A l'opérateur d'inertie ($A\Omega = \boldsymbol{M}$); les vecteurs $\boldsymbol{\Omega}$ et $\boldsymbol{M}$ appartiennent à un système de coordonnées mobile K (§ 26). Le vecteur $\boldsymbol{m} = R\boldsymbol{M}$ du moment cinétique du solide par rapport à O dans l'espace est invariant par le mouvement (§ 28, B).

Donc le vecteur $\boldsymbol{M}$ ($\boldsymbol{M} \in K$) doit se déplacer de telle sorte que le vecteur $\boldsymbol{m} = R_t\boldsymbol{M}(t)$ ne soit pas fonction du temps t.

Théorème. *On a la relation*

$$\frac{d\boldsymbol{M}}{dt} = [\boldsymbol{M}, \Omega]. \qquad (1)$$

D é m o n s t r a t i o n. Appliquons la formule (5) du § 26 à la vitesse du mouvement d'un « point » $\boldsymbol{M}(t) \in K$ par rapport à l'espace fixe k:

$$\dot{\boldsymbol{m}} = R\dot{\boldsymbol{M}} + [\omega, \boldsymbol{m}] = R[\dot{\boldsymbol{M}} + [\Omega, \boldsymbol{M}]).$$

Comme le moment $\boldsymbol{m}$ par rapport à l'espace se conserve ($\dot{\boldsymbol{m}} = 0$), il vient $\dot{\boldsymbol{M}} + [\Omega, \boldsymbol{M}] = 0$, c.q.f.d.

La relation (1) est appelée *équation d'Euler*. Comme $\boldsymbol{M} = A\Omega$, on peut considérer (1) comme une équation différentielle en $\boldsymbol{M}$ (ou en Ω). Si

$$\Omega = \Omega_1 e_1 + \Omega_2 e_2 + \Omega_3 e_3, \quad \boldsymbol{M} = M_1 e_1 + M_2 e_2 + M_3 e_3$$

sont les décompositions de Ω et $\boldsymbol{M}$ suivant les axes d'inertie en O, alors $M_i = I_i \Omega_i$ et (1) prend la forme du système de trois équations

$$\frac{dM_1}{dt} = a_1 M_2 M_3, \quad \frac{dM_2}{dt} = a_2 M_3 M_1, \quad \frac{dM_3}{dt} = a_3 M_1 M_2, \tag{2}$$

où $a_1 = \frac{I_2 - I_3}{I_2 I_3}$, $a_2 = \frac{I_3 - I_1}{I_3 I_1}$, $a_3 = \frac{I_1 - I_2}{I_1 I_2}$, ou bien la forme du système de trois équations en les trois composantes du vecteur vitesse angulaire

$$I_1 \frac{d\Omega_1}{dt} = (I_2 - I_3)\,\Omega_2\Omega_3,$$

$$I_2 \frac{d\Omega_2}{dt} = (I_3 - I_1)\,\Omega_3\Omega_1,$$

$$I_3 \frac{d\Omega_3}{dt} = (I_1 - I_2)\,\Omega_1\Omega_2.$$

R e m a r q u e. Si le solide est soumis à l'action de forces extérieures dont la somme des moments par rapport à O est égale à $\boldsymbol{n}$ dans le système de coordonnées fixe et à $\boldsymbol{N}$ dans le système de coordonnées mobile ($\boldsymbol{n} = R\boldsymbol{N}$), alors

$$\dot{\boldsymbol{m}} = \boldsymbol{n}$$

et l'équation d'Euler s'écrit

$$\frac{d\boldsymbol{M}}{dt} = [\boldsymbol{M}, \Omega[+ \boldsymbol{N}.$$

B. Discussion des solutions de l'équation d'Euler.

Lemme. *L'équation d'Euler* (2) *possède deux intégrales premières quadratiques*

$$2E = \frac{M_1^2}{I_1} + \frac{M_2^2}{I_2} + \frac{M_3^2}{I_3} \quad \text{et} \quad M^2 = M_1^2 + M_2^2 + M_3^2.$$

D é m o n s t r a t i o n. E est invariant en vertu de la loi de conservation de l'énergie, M^2 également en vertu de la loi de conser-

vation du moment $\boldsymbol{m}$, puisque $\boldsymbol{m}^2=\boldsymbol{M}^2=M^2$. Le lemme est démontré.

Donc $\boldsymbol{M}$ est situé sur l'intersection d'un ellipsoïde avec une sphère. Etudions l'allure des courbes d'intersection ; pour cela fixons l'ellipsoïde $E>0$ et faisons varier le rayon M de la sphère (fig. 121).

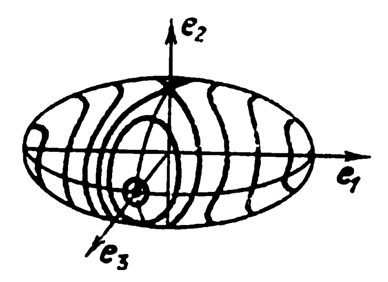


Fig. 121. Trajectoires des équations d'Euler sur une surface de niveau d'énergie.

Supposons pour fixer les idées que $I_1>I_2>I_3$. Les demi-axes de l'ellipsoïde seront $\sqrt{2EI_1}>\sqrt{2EI_2}>\sqrt{2EI_3}$. Si le rayon M de la sphère est inférieur au demi petit axe ou supérieur au demi grand axe ($M<\sqrt{2EI_3}$ ou $M>\sqrt{2EI_1}$), alors l'intersection est vide et aucun mouvement réel ne correspond à ces valeurs de E et M. Si le rayon de la sphère est égal au demi petit axe, alors l'intersection est composée de deux points. Lorsque le rayon augmente ($\sqrt{2EI_3}<$ $<M<\sqrt{2EI_2}$), on obtient deux courbes autour des extrémités du demi petit axe. De la même manière, si le rayon de la sphère est égal au demi grand axe, on obtient les deux extrémités du demi grand axe et s'il est légèrement inférieur, deux courbes fermées au voisinage de ces extrémités. Enfin, si $M=\sqrt{2EI_2}$, l'intersection est constituée de deux cercles.

Chacune des six extrémités des demi-axes de l'ellipsoïde est une trajectoire des équations d'Euler (2) : une position stationnaire du vecteur $\boldsymbol{M}$. Il lui correspond une valeur constante du vecteur vitesse angulaire qui a même support et même sens que l'un des axes d'inertie $\boldsymbol{e}_i$; de plus $\boldsymbol{\Omega}$ et $\boldsymbol{M}$ restent tout le temps colinéaires. Donc le vecteur vitesse angulaire reste colinéaire à $\boldsymbol{m}$: le solide tourne simplement à une vitesse angulaire constante autour d'un axe d'inertie $\boldsymbol{e}_i$ fixe par rapport à l'espace.

D é f i n i t i o n. On appelle *rotation stationnaire* un mouvement tel que la vitesse angulaire du solide soit constante ($\boldsymbol{\omega}=$ const, $\boldsymbol{\Omega}=$ const).

On a démontré le

Théorème. *Un solide ayant un point fixe O admet une rotation stationnaire autour de l'un quelconque de ses trois axes d'inertie* $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$.

Si, comme nous l'avons supposé, $I_1>I_2>I_3$, alors le second membre de l'équation d'Euler ne s'annule nulle part ailleurs, i.e. il n'existe pas d'autres rotations stationnaires.

Etudions maintenant la stabilité des solutions stationnaires de l'équation d'Euler (au sens de Liapounov).

Théorème. *Les solutions stationnaires* $\boldsymbol{M} = M_1e_1$ *et* $\boldsymbol{M} = M_3e_3$ *de l'équation d'Euler correspondant au grand et au petit axe d'inertie sont stables, la solution correspondant à l'axe moyen* ($\boldsymbol{M} = M_2e_2$) *est instable.*

En effet, la trajectoire sera une petite courbe fermée si la condition initiale s'écarte faiblement de M_1e_1 ou M_3e_3, et une grande si elle s'écarte faiblement de M_2e_2.

E x e r c i c e. Dire si les *rotations* stationnaires *d'un solide* autour du grand et du petit axe d'inertie sont stables au sens de Liapounov.

Réponse. Non.

C. Equation de mouvement de Poinsot. Nous savons parfaitement que le mouvement des vecteurs moment et vitesse angulaire par rapport au solide ($\boldsymbol{M}$ et Ω) est périodique si $M \neq \sqrt{2EI_i}$.

Voyons comment tourne le solide *dans l'espace*; pour cela considérons son ellipsoïde d'inertie

$$\mathrm{E} = \{\Omega : (A\Omega, \ \Omega) = 1\} \subset K,$$

où $A : \Omega \to \boldsymbol{M}$ est l'opérateur symétrique d'inertie du solide fixé en O.

A chaque instant, l'ellipsoïde E occupe une position $R_t\mathrm{E}$ par rapport à l'espace immobile k.

Théorème (de Poinsot). *L'ellipsoïde d'inertie roule sans glisser sur un plan fixe perpendiculaire au vecteur moment* $\boldsymbol{m}$ (fig. 122).

D é m o n s t r a t i o n. Considérons un plan π perpendiculaire au vecteur moment $\boldsymbol{m}$ et tangent à l'ellipsoïde d'inertie $R_t\mathrm{E}$. Ces

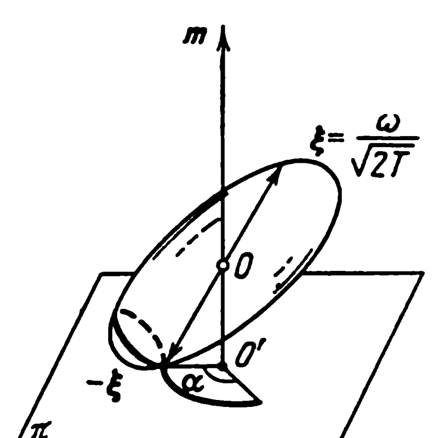


Fig. 122. Roulement de l'ellipsoïde d'inertie sur un plan fixe.

plans sont au nombre de deux et au point de tangence la normale à l'ellipsoïde est parallèle à $\boldsymbol{m}$.

Or au point Ω l'ellipsoïde d'inertie E admet pour normale grad $(A\Omega, \ \Omega) = 2A\Omega = 2\boldsymbol{M}$. Donc aux points $\pm\ \xi = \frac{\omega}{\sqrt{2T}}$ d'intersection

de l'axe ω avec R_tE la normale à R_tE est précisément colinéaire à $\boldsymbol{m}$.

Donc le plan π est tangent à R_tE aux points $\pm\xi$ de l'axe instantané de rotation. Or le produit scalaire de ξ par le vecteur fixe $\boldsymbol{m}$ est égal à $\pm \frac{1}{\sqrt{2T}} (\boldsymbol{m}, \omega) = \pm \sqrt{2T}$, i.e. il est constant. La distance du plan π au point O est constante, i.e. le plan π est fixe.

Comme le point de tangence est situé sur l'axe instantané de rotation, sa vitesse est nulle. Donc l'ellipsoïde R_tE roule sur π sans glisser, c. q. f. d.

Corollaire. *Si les conditions initiales sont proches d'une rotation stationnaire autour du grand (ou du petit) axe d'inertie le vecteur vitesse angulaire reste constamment proche de sa position initiale non seulement par rapport au solide* (Ω) *mais aussi par rapport à l'espace* (ω).

Considérons maintenant la trajectoire du point de tangence sur le plan fixe π. Lorsque le point effectue un tour sur l'ellipsoïde, les conditions initiales se répètent avec la seule différence que le solide tourne d'un certain angle α autour de l'axe $\boldsymbol{m}$. Le deuxième tour sera exactement le même que le premier; si $\alpha = 2\pi \frac{p}{q}$, le mouvement est dans son ensemble périodique; si l'angle est incommensurable avec 2π, le solide ne reviendra jamais à sa position initiale.

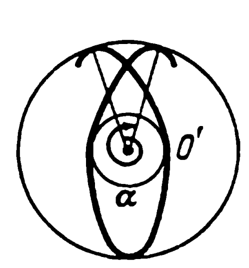


Fig. 123. Trajectoire d'un point de tangence sur un plan fixe.

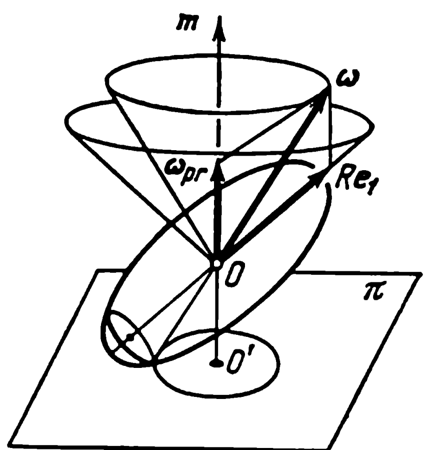


Fig. 124. Roulement d'un ellipsoïde de révolution sur un plan fixe.

Le point de tangence balaie sur le plan π une couronne partout dense de centre O' (fig. 123).

E x e r c i c e. Montrer que les composantes connexes des variétés invariantes V_c de dimension 2 (§ 28, B) dans l'espace $TSO(3)$ de dimension six sont des tores sur lesquels on peut choisir des coordonnées φ_1, φ_2 mod 2π telles que $\dot{\varphi}_1 = \omega_1(c)$, $\dot{\varphi}_2 = \omega_2(c)$.

I n d i c a t i o n. Prendre pour φ_1 la phase de variation périodique de M.

Considérons le cas particulier important où l'ellipsoïde d'inertie est un ellipsoïde de révolution:

$$I_2 = I_3 \neq I_1.$$

Dans ce cas l'axe de l'ellipsoïde, $R_t e_1$, l'axe instantané de rotation ω et le vecteur m sont toujours coplanaires. Les angles qu'ils forment entre eux et le module de ω se conservent; le point de tangence décrit des cercles sur l'ellipsoïde ainsi que sur le plan; les axes de rotation (ω) et de symétrie ($R_t e_1$) décrivent avec la même vitesse angulaire des cônes autour du vecteur moment m (fig. 124).

Ce mouvement de rotation autour de m s'appelle *précession*.

E x e r c i c e. Trouver la vitesse angulaire de précession.

Réponse. Décomposons le vecteur vitesse angulaire ω suivant le vecteur moment m et l'axe $R_t e_1$ du solide. La première composante est justement la vitesse angulaire de précession $\omega_{\text{préc.}} = M/I_2$.

I n d i c a t i o n. Représenter le mouvement du solide sous forme de produit d'une rotation autour de l'axe du moment par une rotation autour de l'axe du solide. La vitesse angulaire du produit des deux mouvements est égale à la somme géométrique des vecteurs de leurs vitesses angulaires.

R e m a r q u e. Un solide ayant un point fixe O, en l'absence de forces extérieures, est un système de Lagrange dont l'espace de configuration est un groupe, plus exactement le groupe SO (3), et dont le lagrangien est invariant par les translations à gauche.

On montre qu'une grande partie de la théorie eulérienne du solide n'utilise que cette circonstance et vaut par conséquent pour tout système lagrangien invariant à gauche sur un groupe de Lie arbitraire.

En particulier, en appliquant cette théorie à un groupe de difféomorphismes d'un domaine riemannien D conservant l'élément de volume on déduirait les principaux théorèmes de la dynamique des fluides parfaits.

§ 30. Toupie de Lagrange

On étudie le mouvement d'un solide à symétrie axiale, ayant un point fixe, dans un champ de forces homogène. Ce mouvement est composé de trois mouvements périodiques: une rotation, une précession et une nutation.

A. Angles d'Euler. Considérons un solide ayant un point fixe O soumis à l'action de la force de pesanteur mg. Le mouvement d'un tel « solide pesant » est un problème qui n'a encore pas été résolu dans le cas général et il est même irrésoluble dans un certain sens.

Dans ce problème à trois degrés de liberté on ne connaît que deux intégrales premières: l'énergie $E = T + U$ et la projection M_z du moment cinétique sur la verticale.

Il existe un cas particulier important où ce problème se résout complètement, c'est le cas de la *toupie symétrique*. Une toupie symétrique ou lagrangienne est par définition un solide ayant un point

fixe O, dont l'ellipsoïde d'inertie en O est un ellipsoïde de révolution et dont le centre de gravité est situé sur l'axe de rotation e_3 (fig. 125).

Dans ce cas la fonction de Lagrange est invariante par toute rotation autour de e_3 et en vertu du théorème de Noether il doit exister en plus de E et M_z une troisième intégrale première (nous verrons que c'est la projection M_3 du vecteur moment sur l'axe e_3).

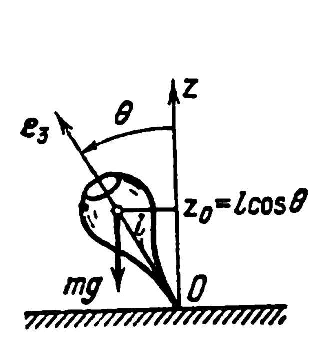




Fig. 125. Toupie de Lagrange.

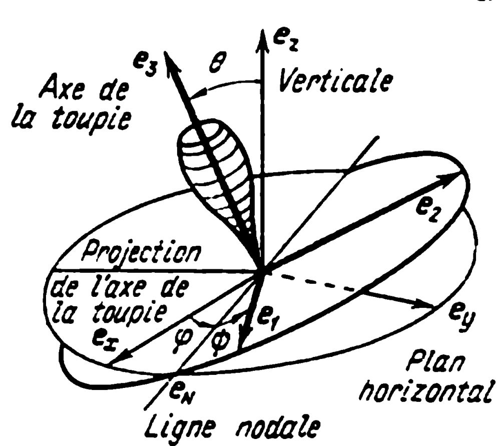


Fig. 126. Angles d'Euler.

Si l'on réussit à introduire trois coordonnées telles qu'elles comprennent les angles de rotation autour de l'axe z et autour de l'axe de la toupie, alors ces coordonnées seront cycliques et le problème à trois degrés de liberté se ramènera à un problème à un degré de liberté (pour la troisième coordonnée).

Dans l'espace de configuration SO(3), un tel choix est possible: ces coordonnées φ, ψ, θ s'appellent *angles d'Euler* et forment dans SO(3) un système de coordonnées local semblable aux coordonnées géographiques de la sphère, i. e. avec des singularités aux pôles et une multivocité sur un méridien.

Introduisons les *notations* suivantes (fig. 126):

e_x, e_y, e_z les vecteurs unitaires d'un repère fixe orthonormé de sens direct et d'origine O;

e_1, e_2, e_3 les vecteurs unitaires d'un repère orthonormé de sens direct lié au solide et orienté suivant les axes d'inertie du solide en O;

$I_1 = I_2 \neq I_3$ les moments d'inertie du solide au point O;

e_N le vecteur unitaire de l'axe $[e_z, e_3]$ appelé ligne nodale. (Tous les vecteurs sont contenus dans l'« espace fixe » k.)

Pour transférer le repère fixe (e_x, e_y, e_z) dans le repère mobile (e_1, e_2, e_3) il faut effectuer trois rotations:

1) *une rotation d'angle* φ *autour de l'axe* e_z. L'axe e_z reste en place et l'axe e_x vient en e_N;

2) *une rotation d'angle* θ *autour de l'axe* e_N qui laisse e_N en place et amène e_z sur e_3;

3) et enfin *une rotation d'angle* ψ *autour de* e_3 qui transporte e_N sur e_1 et laisse e_3 en place.

Au terme de ces trois rotations e_x vient en e_1 et e_z en e_3, donc e_y coïncidera avec e_2.

Les angles φ, ψ, θ sont appelés *angles d'Euler*. On démontre sans peine le

Théorème. *A tout triplet* φ, θ, ψ *correspond une rotation* $R(\varphi, \theta, \psi) \in \in SO(3)$ *de l'espace transférant le repère* (e_x, e_y, e_z) *dans le repère* (e_1, e_2, e_3). *Ceci étant, l'application* $(\varphi, \theta, \psi) \to R(\varphi, \theta, \psi)$ *définit des coordonnées locales*

$$0 < \varphi < 2\pi, \quad 0 < \psi < 2\pi, \quad 0 < \theta < \pi$$

sur l'espace de configuration SO(3) *de la toupie.*

Comme dans le cas de la longitude, φ et ψ peuvent être considérés comme des angles de mod 2π; pour $\theta = 0$ ou $\theta = \pi$ l'application $(\varphi, \theta, \psi) \to R$ présente une singularité du type pôle.

B. Calcul du lagrangien. Exprimons le lagrangien en fonction des coordonnées φ, θ, ψ et de leurs dérivées.

L'énergie potentielle est visiblement égale à

$$U = \iiint zg\, dm = mg\, z_0 = mgl \cos\theta,$$

où z_0 est la cote du centre de gravité par rapport à O (fig. 125).

Evaluons l'énergie cinétique. On aura besoin d'un petit artifice: considérons le *cas particulier* où $\varphi = \psi = 0$.

Lemme. *La vitesse angulaire de la toupie s'exprime en fonction des dérivées des angles d'Euler par la formule*

$$\omega = \dot\theta e_1 + (\dot\varphi \sin\theta)\, e_2 + (\dot\psi + \dot\varphi \cos\theta)\, e_3$$

si $\varphi = \psi = 0$.

D é m o n s t r a t i o n. Considérons la vitesse d'un point de la toupie occupant la position r à la date t. Au bout d'un temps dt ce point occupera la position (à $(dt)^2$ près)

$$R(\varphi + d\varphi, \theta + d\theta, \psi + d\psi)\, R^{-1}(\varphi, \theta, \psi)\, r,$$

où

$$d\varphi = \dot\varphi\, dt, \quad d\theta = \dot\theta\, dt, \quad d\psi = \dot\psi\, dt.$$

Donc le vecteur déplacement est, avec la même précision, la somme des trois termes

$$R(\varphi + d\varphi, \theta, \psi)\, R^{-1}(\varphi, \theta, \psi)\, r - r = [\omega_\varphi, r]\, dt,$$

$$R(\varphi, \theta + d\theta, \psi)\, R^{-1}(\varphi, \theta, \psi)\, r - r = [\omega_\theta, r]\, dt,$$

$$R(\varphi, \theta, \psi + d\psi)\, R^{-1}(\varphi, \theta, \psi)\, r - r = [\omega_\psi, r]\, dt$$

(les vitesses angulaires ω_φ, ω_θ et ω_ψ sont définies par ces formules).

Donc la vitesse du point $\boldsymbol{r}$ est $\dot{\boldsymbol{r}} = [\omega_\varphi + \omega_\theta + \omega_\psi, \boldsymbol{r}]$, et la vitesse angulaire du solide est

$$\omega = \omega_\varphi + \omega_0 + \omega_\psi,$$

où les termes sont définis par les formules précédentes.

Il reste à décomposer les vecteurs ω_φ, ω_θ et ω_ψ suivant $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. Utilisons maintenant le fait que $\varphi = \psi = 0$. Il en résulte que

$$R(\varphi + d\varphi, \theta, \psi) R^{-1}(\varphi, \theta, \psi)$$

est une simple rotation d'angle $d\varphi$ autour de l'axe $\boldsymbol{e}_z$, de sorte que

$$\omega_\psi = \dot{\varphi}\boldsymbol{e}_z.$$

D'autre part, si $\varphi = \psi = 0$, alors $R(\varphi, \theta + d\theta, \psi) R^{-1}(\varphi, \theta, \psi)$ est une simple rotation d'angle $d\theta$ autour de l'axe $\boldsymbol{e}_N = \boldsymbol{e}_x = \boldsymbol{e}_1$, de sorte que

$$\omega_\theta = \dot{\theta}\boldsymbol{e}_1.$$

Enfin, $R(\varphi, \theta, \psi + d\psi) R^{-1}(\varphi, \theta, \psi)$ est une rotation d'angle $d\psi$ autour de l'axe $\boldsymbol{e}_3$, donc

$$\omega_\psi = \dot{\psi}\boldsymbol{e}_3.$$

Finalement, pour $\varphi = \psi = 0$ on obtient

$$\omega = \dot{\varphi}\boldsymbol{e}_z + \dot{\theta}\boldsymbol{e}_1 + \dot{\psi}\boldsymbol{e}_3.$$

Or il est évident que

$$\boldsymbol{e}_z = \boldsymbol{e}_3 \cos\theta + \boldsymbol{e}_2 \sin\theta.$$

Donc les composantes de la vitesse angulaire suivant les axes d'inertie $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ sont

$$\omega_1 = \theta, \quad \omega_2 = \dot{\varphi}\sin\theta, \quad \omega_3 = \dot{\psi} + \dot{\varphi}\cos\theta, \text{ c.q.f.d.}$$

Comme $T = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$, l'énergie cinétique est alors donnée par la formule

$$T = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)^2.$$

Or l'énergie cinétique ne peut dépendre ni de φ ni de ψ : ce sont des coordonnées cycliques et sans changer T on peut toujours faire en sorte que $\varphi = 0$ et $\psi = 0$.

Donc la formule obtenue pour l'énergie cinétique vaut pour tous les φ et ψ.

Nous obtenons donc le lagrangien

$$L = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)^2 - mgl\cos\theta.$$

D. Etude du mouvement. Aux coordonnées cycliques φ, ψ correspondent les intégrales premières

$$\frac{\partial L}{\partial \dot{\varphi}} = M_z = \dot{\varphi}\,(I_1 \sin^2\theta + I_3 \cos^2\theta) + \dot{\psi} I_3 \cos\theta,$$

$$\frac{\partial L}{\partial \dot{\psi}} = M_3 = \dot{\varphi} I_3 \cos\theta + \dot{\psi} I_3.$$

Théorème. *L'angle d'écart* θ *entre l'axe de la toupie et la verticale varie en fonction du temps comme dans un système de dimension* 1 *dont l'énergie serait*

$$E' = \frac{I_1}{2}\dot{\theta}^2 + U_{\text{eff}}(\theta),$$

l'énergie potentielle effective U_{eff} *étant donnée par la formule*

$$U_{\text{eff}} = \frac{(M_z - M_3\cos\theta)^2}{2I_1\sin^2\theta} + mgl\cos\theta.$$

D é m o n s t r a t i o n. Exprimons $\dot{\varphi}$ et $\dot{\psi}$ en fonction de M_3 et M_z en vertu de la théorie générale. L'énergie totale du système prend la forme

$$E = \frac{I_1}{2}\dot{\theta}^2 + \frac{M_3^2}{2I_3} + mgl\cos\theta + \frac{(M_z - M_3\cos\theta)^2}{2I_1\sin^2\theta}$$

et

$$\dot{\varphi} = \frac{M_z - M_3\cos\theta}{I_1\sin^2\theta}.$$

Le terme $\frac{M_3^2}{2I_3} = E - E'$ ne contenant pas θ n'influe pas sur l'équation en θ. Donc le théorème est démontré.

Etudions maintenant le système de dimension 1 obtenu. Pour la commodité faisons le changement $\cos\theta = u$ $(-1 \leqslant u \leqslant 1)$.

En adoptant d'autre part les notations

$$\frac{M_z}{I_1} = a, \quad \frac{M_3}{I_1} = b, \quad \frac{2E'}{I_1} = \alpha, \quad \frac{2mgl}{I_1} = \beta > 0,$$

on peut écrire la loi de conservation de l'énergie E' sous la forme

$$\dot{u}^2 = f(u), \text{ où } f(u) = (\alpha - \beta u)(1 - u^2) - (a - bu)^2,$$

et la loi de variation de l'azimut φ sous la forme

$$\dot{\varphi} = \frac{a - bu}{1 - u^2}.$$

Remarquons que $f(u)$ est un polynôme du troisième degré, $f(+\infty) = +\infty$ et $f(\pm 1) = -(a \mp b)^2 < 0$ si seulement $a \neq \pm b$. Par ailleurs, à un mouvement réel ne sont associées que des valeurs constantes a, b, α, β telles que $f(u) \geqslant 0$ pour certains $-1 \leqslant u \leqslant 1$.

Donc $f(u)$ possède exactement deux racines réelles u_1 et u_2 comprises dans l'intervalle $[-1, 1]$ (et une pour $u > 1$, fig. 127).

Fig. 127. Graphe de la fonction $f(u)$.

Donc l'angle θ varie de façon périodique entre deux valeurs limites θ_1, θ_2 (fig. 128).

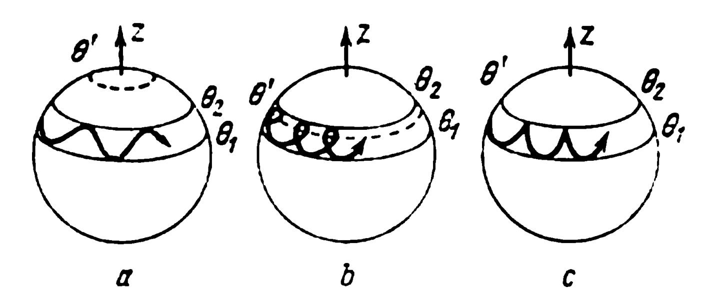


Fig. 128. Trace de l'axe de la toupie sur la sphère unitaire.

Cette variation périodique est appelée *nutation*.

Voyons maintenant le mouvement de l'axe de la toupie en fonction de l'azimut. Le point d'intersection de l'axe avec la sphère unité se déplace sur la zone sphérique comprise entre les parallèles θ_1 et θ_2, les variations de l'azimut de l'axe obéissant à l'équation

$$\dot{\varphi} = \frac{a - bu}{1 - u^2}.$$

Si la racine u' de l'équation $a = bu$ est située à l'extérieur de $]u_1, u_2[$, alors l'angle φ varie de façon monotone et l'axe décrit une courbe de type sinusoïde sur la sphère unité (fig. 128, *a*).

Si la racine u' de l'équation $a = bu$ est comprise dans l'intervalle $]u_1, u_2[$, la vitesse de variation de φ est opposée sur les parallèles θ_1 et θ_2 et l'axe décrit sur la sphère une courbe comportant des boucles (fig. 128, *b*).

Si enfin la racine u' de l'équation $a = bu$ coïncide avec l'une des valeurs u_1 ou u_2 (disons avec u_2), l'axe décrit une courbe présentant des pointes (fig. 128, *c*).

Le dernier cas quoique exceptionnel s'observe chaque fois que l'on lâche l'axe de la toupie qu'on aura mise en mouvement sous un angle θ_2 sans vitesse initiale : la toupie tombe puis se relève.

Le mouvement azimutal de l'axe est appelé *précession*. Le mouvement de la toupie est composé d'une rotation propre, d'une nutation et d'une précession possédant chacune sa propre fréquence. Si

ces fréquences sont incommensurables, la toupie ne revient jamais à son état initial bien qu'elle s'en approche d'une distance aussi petite que l'on veut.

§ 31. Toupie dormante et toupie rapide

Les formules obtenues dans le paragraphe 30 ramènent la résolution des équations de mouvement de la toupie à des quadratures elliptiques. Cependant, en pratique, il est plus facile d'étudier ce mouvement sans recourir aux quadratures.

Dans ce paragraphe on analyse la stabilité de la position verticale de la toupie et on donne les formules approchées du mouvement de la toupie lancée rapidement.

A. Toupie dormante. Considérons d'abord une solution particulière des équations de mouvement, telle que l'axe de la toupie

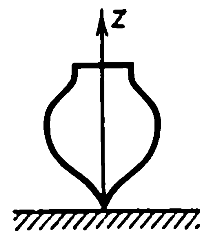


Fig. 129. Toupie dormante.

reste tout le temps vertical ($\theta = 0$) et la vitesse angulaire constante (toupie dormante). De toute évidence on aura $M_z = M_3 = I_3\omega_3$ (fig. 129).

E x e r c i c e. Montrer que la rotation stationnaire autour de la verticale est toujours instable au sens de Liapounov.

Nous étudions le mouvement de *l'axe de la toupie* et non pas de la toupie elle-même. L'axe de la toupie sera-t-il animé d'un mouvement de rotation stable au voisinage de la verticale, i.e. θ restera-t-il petit? Développons l'énergie potentielle effective du système

$$U_{\text{eff}} = \frac{(M_z - M_3 \cos\theta)^2}{2I_1 \sin^2\theta} + mgl\cos\theta$$

en série suivant les puissances de θ:

$$U_{\text{eff}} = \frac{I_3^2\omega_3^2 \frac{\theta^4}{4}}{2I_1\theta^2} + \ldots - mgl\frac{\theta^2}{2} + \ldots = C + A\theta^2 + \ldots,$$

$$A = \frac{\omega_3^2 I_3^2}{8I_1} - \frac{mgl}{2}.$$

Si $A > 0$, la position d'équilibre $\theta = 0$ du système de dimension 1 est stable, et si $A < 0$, instable. Donc la condition de stabilité prend la forme

$$\omega_3^2 > \frac{4mglI_1}{I_3^2}.$$

Lorsque les frottements font baisser la vitesse de la toupie dormante au-dessous de cette limite, la toupie se « réveille ».

E x e r c i c e. Montrer que pour $\omega_3^2 > \frac{4mglI_1}{I_3^2}$ l'axe de la toupie dormante est stable relativement aux perturbations qui modifient les valeurs de M_z, M_3 et pas seulement de θ.

B. Toupie rapide. La toupie est *rapide* si l'énergie cinétique de sa rotation est grande devant son énergie potentielle:

$$\frac{1}{2} I_3\omega_3^2 \gg mgl.$$

Des considérations de similitude entraînent que l'accroissement de la vitesse angulaire de N fois équivaut très exactement à une diminution du poids de N^2 fois.

Théorème. *Si tout en conservant la position initiale de la toupie on accroît de N fois sa vitesse angulaire, sa trajectoire sera exactement la même que si la vitesse angulaire n'avait pas changé mais l'accélération de la pesanteur avait diminué de N^2 fois. Ceci étant, si la vitesse angulaire est grande, la trajectoire est parcourue N fois plus vite* *).

Donc nous pouvons analyser le cas $g \to 0$ et appliquer les résultats obtenus à l'étude du cas $\omega \to \infty$.

Considérons d'abord le cas $g = 0$, i. e. le mouvement d'une toupie symétrique en l'absence de la force de pesanteur. Comparons les deux descriptions de ce mouvement: celle de Lagrange (§ 30, D) et celle de Poinsot (§ 29, C).

Etudions d'abord l'équation de Lagrange pour les variations de l'angle de nutation θ.

Lemme. *En l'absence de la force de pesanteur, l'angle θ_0 tel que $M_z = M_3 \cos\theta_0$ est position d'équilibre stable de l'équation du mouvement de l'axe de la toupie.*

La fréquence des petites oscillations de θ au voisinage de cette position d'équilibre est égale à

$$\omega_{\text{nut}} = \frac{I_3\omega_3}{I_1}.$$

D é m o n s t r a t i o n. En l'absence de la force de pesanteur l'énergie potentielle effective se réduit à

$$U_{\text{eff}} = \frac{(M_z - M_3\cos\theta)^2}{2I_1\sin^2\theta}.$$

Cette fonction non négative possède un minimum nul pour $\theta = \theta_0$ qui est défini à partir de la condition $M_z = M_3 \cos\theta_0$ (fig. 130).

*) Désignons par $\varphi_g(t, \xi)$ la position occupée par la toupie à l'instant t avec la condition initiale $\xi \in TSO(3)$ et l'accélération g de la pesanteur. Le théorème affirme que $\varphi_g(t, N\xi) = \varphi_{N^{-2}g}(Nt, \xi)$.

Donc, l'angle θ_0 que fait l'axe de la toupie avec la verticale est stationnaire stable: si la valeur initiale de θ s'écarte faiblement de θ_0, alors θ effectuera des oscillations périodiques autour de θ_0 (nutation). La fréquence de ces oscillations

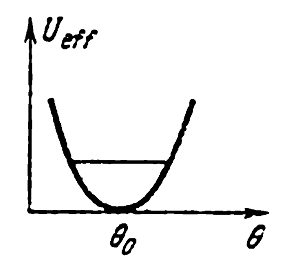


Fig. 130. Energie potentielle effective de la toupie.

se déduit sans peine de la formule générale: la fréquence des petites oscillations d'un système de dimension 1 d'énergie

$$E = \frac{a\dot{x}^2}{2} + U(x), \qquad U(x_0) = \min U(x)$$

est définie par la formule (§ 22, D)

$$\omega^2 = \frac{U''(x_0)}{a}.$$

L'énergie du système de dimension 1 décrivant les oscillations de θ s'écrit

$$\frac{I_1}{2}\dot{\theta}^2 + U_{eff}.$$

Pour $\theta = \theta_0 + x$ il vient $M_z - M_3 \cos\theta = M_3(\cos\theta_0 - \cos(\theta_0 + x)) = M_3 x \sin\theta_0 + O(x^2)$

$$u_{eff} = \frac{M_3^2 \cdot x^2 \cdot \sin^2\theta_0}{2I_1 \sin^2\theta_0} + o(x^2) = \frac{I_3^2\omega_3^2}{2I_1} x^2 + \ldots,$$

d'où l'on déduit l'expression de la fréquence de nutation

$$\omega_{nut} = \frac{I_3\omega_3}{I_1}, \quad \text{c. q. f. d.}$$

De la formule $\dot{\varphi} = \frac{M_z - M_3\cos\theta}{I_1 \sin^2\theta}$ on déduit que pour $\theta = \theta_0$ l'azimut de l'axe n'est pas fonction du temps: l'axe est fixe. On aurait pu également se servir de cette formule pour étudier le mouvement azimutal de l'axe pour de faibles écarts $\theta - \theta_0$. Mais nous allons agir autrement.

En l'absence de la force de pesanteur, le mouvement de la toupie peut être considéré comme un mouvement au sens de Poinsot. Donc l'axe de la toupie est en rotation uniforme autour du vecteur moment cinétique qui est fixe.

Ainsi donc l'axe de la toupie décrit sur la sphère unité un cercle dont le centre correspond au vecteur moment cinétique (fig. 131).

R e m a r q u e. En résumé, le mouvement de *nutation* de Lagrange n'est autre que le mouvement de *précession* de Poinsot.

Il est évident que la formule obtenue plus haut pour la fréquence de la petite nutation $\omega_{\text{nut}} = I_3\omega_3/I_1$ concorde avec la fréquence de

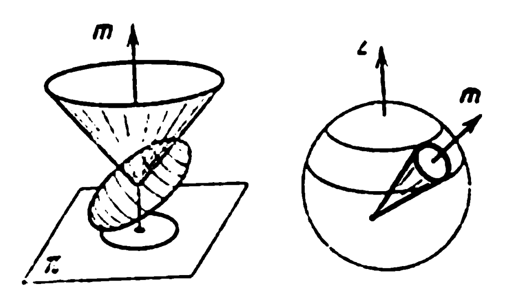


Fig. 131. Comparaison des mouvements de Lagrange et de Poinsot de la toupie.

la précession $\omega = M/I_1$ dans le mouvement de Poinsot: lorsque l'amplitude de la nutation tend vers 0, alors $I_3\omega_3 \to M$.

C. Toupie dans un faible champ. Voyons maintenant le cas où la force de pesanteur existe, mais est de faible intensité (M_z et M_3 sont fixes). Dans ce cas à l'énergie potentielle effective on ajoute le terme $mgl \cos\theta$ qui est petit avec ses dérivées. Montrons que ce terme ne modifie pratiquement pas la fréquence de nutation.

Lemme. *Supposons qu'une fonction $f(x)$ admette un minimum pour $x = 0$ et le développement de Taylor $f(x) = A\frac{x^2}{2} + \ldots$, $A > 0$. Soit $h(x)$ une fonction admettant en 0 le développement taylorien $h(x) = B + Cx + \ldots$. Alors*

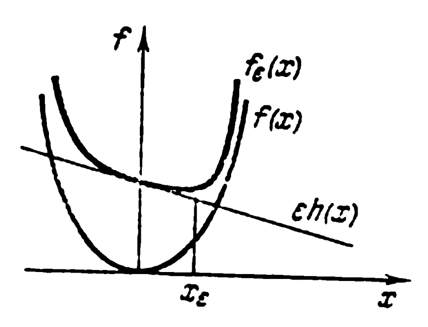


Fig. 132. Déplacement du minimum par une faible variation de la fonction.

pour ε suffisamment petit $f_\varepsilon(x) = f(x) + \varepsilon h(x)$ présente un minimum en un point (fig. 132)

$$x_\varepsilon = -\frac{C\varepsilon}{A} + O(\varepsilon^2)$$

proche de 0 et en outre $f_\varepsilon''(x_l) = A + O(\varepsilon)$.

En effet, on a $f_\varepsilon'(x) = Ax + C\varepsilon + O(x^2)$ et le résultat annoncé s'obtient par application du théorème des fonctions implicites à $f_\varepsilon'(x)$.

En vertu du lemme, l'énergie potentielle effective possède un point de minimum θ_g proche de θ_0 pour les petits g, et en outre $U''(\theta_g) \approx U''(\theta_0)$. Donc la fréquence de la petite nutation au voisinage de θ_0 est proche de la valeur obtenue pour $g = 0$:

$$\lim_{g\to 0} \omega_{\text{nut}} = \frac{I_3}{I_1}\,\omega_3.$$

E. Toupie lancée rapidement. Considérons maintenant des conditions initiales spéciales, plus exactement nous relâchons l'axe de la toupie sans poussée initiale à partir de l'angle θ_0.

Théorème. *Si à l'instant initial l'axe de la toupie est fixe* ($\dot{\varphi} = \dot{\theta} = 0$) *et la toupie tourne rapidement autour de son axe* ($\omega_3 \to \infty$) *avec un angle de nutation* θ_0 ($M_z = M_3 \cos\theta_0$), *alors asymptotiquement pour* $\omega_3 \to \infty$:

1) *la fréquence de nutation est proportionnelle à la vitesse angulaire*;

2) *l'amplitude de nutation est inversement proportionnelle au carré de la vitesse angulaire*;

3) *la fréquence de précession est inversement proportionnelle à la vitesse angulaire*;

4) *sont valables les formules asymptotiques* (*pour* $\omega_3 \to \infty$)

$$\omega_{\text{nut}} \sim \frac{I_3}{I_1}\omega_3, \quad a_{\text{nut}} \sim \frac{I_1 mgl}{I_3^2\omega_3^2}\sin\theta_0, \quad \omega_{\text{préc}} \sim \frac{mgl}{I_3\omega_3}$$

$\left(\textit{ici } f(\omega_3) \sim g(\omega_3), \textit{ si } \lim\limits_{\omega_3\to\infty} \frac{f}{g} = 1\right)$.

Pour démontrer ce théorème passons au cas où la vitesse angulaire initiale est fixe, mais $g \to 0$.

En interprétant ensuite les formules obtenues à l'aide de la similitude (voir point B) on obtient le théorème formulé.

Nous savons déjà (cf. § 30, E) que sous ces conditions initiales l'axe de la toupie décrit sur la sphère unité une courbe présentant des pointes.

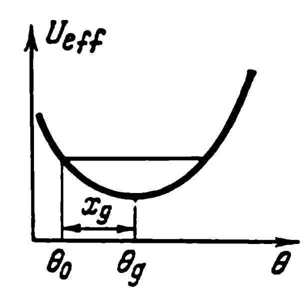


Fig. 133. Détermination de l'amplitude de nutation.

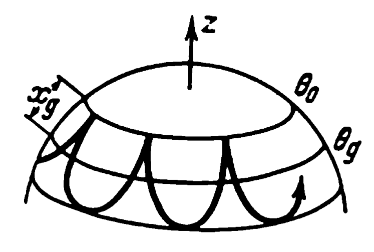


Fig. 134. Mouvement de l'axe de la toupie.

Appliquons le lemme à la recherche du point de minimum de l'énergie potentielle effective. Posons (fig. 133)

$$\theta = \theta_0 + x, \quad \cos\theta = \cos\theta_0 - x\sin\theta_0 + \ldots .$$

Comme plus haut nous obtenons les développements tayloriens suivant les puissances de x:

$$U_{\text{eff}|g=0} = \frac{I_3^2\omega_3^2}{2I_1}x^2 + \ldots, \quad mgl\cos\theta = mgl\cos\theta_0 - xmgl\sin\theta_0 + \ldots .$$

Appliquant le lemme à $f = U_{\text{eff}|g=0}$, $g = \varepsilon$, $h = ml\cos(\theta_0 + x)$ on constate que le minimum de l'énergie potentielle effective U_{eff} est atteint pour

$$\theta_g = \theta_0 + x_g, \quad x_g = \frac{I_1 ml \sin\theta_0}{I_3^2\omega_3^2} g + O(g^2).$$

Donc l'angle de nutation θ oscillera autour de θ_g (fig. 134). Or, à l'instant initial $\theta = \theta_0$ et $\dot\theta = 0$. Donc θ_0 correspond à la plus haute position de l'axe de la toupie. Pour les petits g l'amplitude de la nutation est asymptotiquement égale à

$$a_{\text{nut}} \sim x_g \sim \frac{I_1 ml \sin\theta_0}{I_3^2\omega_3^2} g \quad (g \longrightarrow 0).$$

Calculons maintenant la précession de l'axe. De la formule générale

$$\dot\varphi = \frac{M_z - M_3\cos\theta}{I_1\sin^2\theta}$$

pour $M_z = M_3\cos\theta_0$, $\theta = \theta_0 + x$ on trouve $M_z - M_3\cos\theta = M_3 x\sin\theta_0 + {}$ $+ \ldots$, donc

$$\dot\varphi = \frac{M_3}{I_1\sin\theta_0} x + \ldots .$$

Or x oscille harmoniquement entre 0 et $2x_g$ (avec une précision de $O(g^2)$). Donc au cours d'une période de nutation, la valeur moyenne de la vitesse de précession est asymptotiquement égale à

$$\bar{\dot\varphi} \sim \frac{M_3}{I_1\sin\theta_0} x_g \sim \frac{mgl}{I_3\omega_3} \quad (g \longrightarrow 0).$$

Exercice. Montrer que

$$\lim_{g\to 0}\lim_{t\to\infty} \frac{\dfrac{\varphi(t) - \varphi(0)}{t}}{\dfrac{mgl}{I_3\omega_3}} = 1.$$

TROISIÈME PARTIE

MÉCANIQUE DE HAMILTON

La mécanique hamiltonienne est une géométrie dans l'espace des phases. L'espace des phases possède une structure de variété symplectique. Un groupe de difféomorphismes symplectiques évolue sur une variété symplectique. Les principales notions et théorèmes de la mécanique hamiltonienne (même s'ils sont formulés en termes de coordonnées locales symplectiques) sont invariants par ce groupe (et par un groupe plus large de transformations qui opère également sur le temps).

Un système mécanique hamiltonien est défini par une variété de dimension paire (« espace des phases »), une structure symplectique sur cette variété (« invariant intégral de Poincaré ») et une fonction sur cette variété (« fonction de Hamilton »). Tout groupe à un paramètre de difféomorphismes symplectiques de l'espace des phases laissant invariante la fonction de Hamilton est lié à une intégrale première des équations de mouvement.

La mécanique de Lagrange est un cas particulier de celle de Hamilton (l'espace des phases est le fibré cotangent à l'espace de configuration, et la fonction de Hamilton, la transformée de Legendre du lagrangien).

Le point de vue hamiltonien permet d'étudier complètement de nombreux problèmes insolubles par les autres méthodes (par exemple le problème de l'attraction de deux centres fixes, le problème des géodésiques sur un ellipsoïde à trois axes). Le point de vue hamiltonien est encore plus important dans les méthodes approchées de la théorie des perturbations (mécanique céleste), pour la compréhension du caractère général du mouvement de systèmes mécaniques complexes (théorie ergodique, mécanique statistique) ainsi que dans d'autres domaines de la physique mathématique (optique, mécanique quantique, etc.).

CHAPITRE 7

FORMES DIFFÉRENTIELLES

On rencontre les formes différentielles extérieures lorsqu'on généralise au cas multidimensionnel des notions telles que le travail d'un champ pour un déplacement et le flux d'un liquide à travers une surface.

La mécanique hamiltonienne est inconcevable sans l'étude des formes différentielles. Les notions sur les formes différentielles qui nous seront utiles sont le produit extérieur, la différentiation extérieure, l'intégration et la formule de Stokes.

§ 32. Formes extérieures

On définit ici les formes algébriques extérieures.

A. 1-formes. Soit $\mathbf{R}^n$ un espace vectoriel réel de dimension n *). Les vecteurs de cet espace seront notés $\xi, \eta, \ldots$.

Définition. On appelle *forme de degré* 1 ou simplement 1-*forme* une fonction linéaire d'un vecteur, $\omega : \mathbf{R}^n \to \mathbf{R}$,

$$\omega (\lambda_1 \xi_1 + \lambda_2 \xi_2) = \lambda_1 \omega (\xi_1) + \lambda_2 \omega (\xi)_2, \quad \forall \lambda_1, \ \lambda_2 \in \mathbf{R}, \quad \xi_1, \ \ \xi_2 \in \mathbf{R}^n.$$

Rappelons les principaux résultats que l'on connaît sur les 1-formes du cours d'algèbre linéaire. L'ensemble de toutes les 1-formes se transforme en espace vectoriel réel si on définit la somme de formes par la formule

$$(\omega_1 + \omega_2) (\xi) = \omega_1 (\xi) + \omega_2 (\xi),$$

et le produit d'une forme par un nombre par la formule

$$(\lambda \omega) (\xi) = \lambda \omega (\xi).$$

L'espace des 1-formes sur $\mathbf{R}^n$ est lui aussi de dimension n et s'appelle également espace dual $(\mathbf{R}^n)^*$.

Supposons que dans $\mathbf{R}^n$ on ait choisi un système de coordonnées linéaire $x_1, \ldots, x_n$. Toute composante, par exemple x_i, est une

*) Il est essentiel de remarquer que nous n'avons fixé dans $\mathbf{R}^n$ aucune structure euclidienne spéciale. Dans certains *exemples* on utilise une telle structure : on dira alors: « espace euclidien $\mathbf{R}^n$ ».

1-forme. Ces n 1-formes sont linéairement indépendantes. Aussi chaque 1-forme ω a pour expression

$$\omega = a_1x_1 + \ldots + a_nx_n, \quad a_i \in \mathbf{R}.$$

La valeur de ω sur un vecteur ξ est égale à

$$\omega(\xi) = a_1x_1(\xi) + \ldots + a_nx_n(\xi),$$

où $x_1(\xi), \ldots, x_n(\xi)$ sont les composantes du vecteur ξ dans le système de coordonnées choisi.

E x e m p l e. Si dans l'espace euclidien $\mathbf{R}^3$ est donné un champ de forces

Fig. 135. Le travail d'une force est une 1-forme du déplacement.

F(force)
$\omega(\xi)=(F,\xi)$
ξ(déplacement)

homogène F, alors le travail A pour un déplacement ξ est une 1-forme de ξ (fig. 135).

B. 2-formes.

D é f i n i t i o n. On appelle *forme extérieure de degré* 2 ou simplement 2-*forme* une fonction d'un couple de vecteurs, $\omega^2 : \mathbf{R}^n \times \times \mathbf{R}^n \to \mathbf{R}$ bilinéaire et antisymétrique :

$$\omega^2(\lambda_1\xi_1 + \lambda_2\xi_2, \xi_3) = \lambda_1\omega^2(\xi_1, \xi_3) + \lambda_2\omega^2(\xi_2, \xi_3),$$
$$\omega^2(\xi_1, \xi_2) = -\omega^2(\xi_2, \xi_1),$$
$$\forall \lambda_1, \lambda_2 \in \mathbf{R}, \quad \xi_1, \xi_2 \in \mathbf{R}^n.$$

E x e m p l e 1. Soit $S(\xi_1, \xi_2)$ l'aire orientée du parallélogramme construit sur les vecteurs ξ_1, ξ_2 du plan euclidien orienté $\mathbf{R}^2$, i.e.

$$S(\xi_1, \xi_2) = \begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{vmatrix}, \quad \text{où} \quad \begin{matrix} \xi_1 = \xi_{11}e_1 + \xi_{12}e_2, \\ \xi_2 = \xi_{21}e_1 + \xi_{22}e_2, \end{matrix}$$

où e_1, e_2 est un repère définissant l'orientation de $\mathbf{R}^2$.

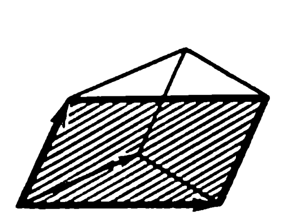


Fig. 136. L'aire orientée est une 2-forme.

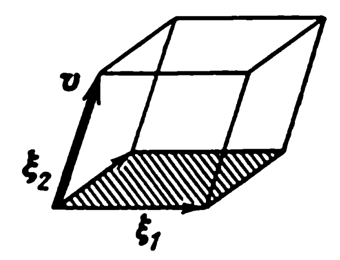


Fig. 137. Le flux de fluide à travers une aire est une 2-forme.

On voit aisément que $S(\xi_1, \xi_2)$ est une 2-forme (fig. 136).

E x e m p l e 2. Soit v un champ de vitesses homogène d'un fluide dans l'espace euclidien orienté $\mathbf{R}^3$ (fig. 137). Le flux du fluide à travers l'aire du

parallélogramme ξ_1, ξ_2 est une fonction bilinéaire antisymétrique de ξ_1, ξ_2, i.e. la 2-forme

$$\omega^2(\xi_1, \xi_2) = (v, \xi_1, \xi_2).$$

Exemple 3. L'aire orientée de la projection d'un parallélogramme de côtés ξ_1, ξ_2 de l'espace euclidien R^3 sur le plan (x_1, x_2) est une 2-forme.

Exercice 1. Montrer que toute 2-forme ω^2 sur R^n vérifie

$$\omega^2(\xi, \xi) = 0, \quad \forall \xi \in R^n.$$

Solution. L'antisymétrie entraîne $\omega^2(\xi, \xi) = -\omega^2(\xi, \xi)$.

L'ensemble de toutes les 2-formes sur R^n forme un espace vectoriel réel si l'on y définit la somme des formes comme

$$(\omega_1 + \omega_2)(\xi_1, \xi_2) = \omega_1(\xi_1, \xi_2) + \omega_2(\xi_1, \xi_2),$$

et le produit d'une 2-forme par un nombre comme

$$(\lambda\omega)(\xi_1, \xi_2) = \lambda\omega(\xi_1, \xi_2).$$

Exercice 2. Montrer que cet espace est de dimension finie. Trouver cette dimension.

Réponse. $\dfrac{n(n-1)}{2}$; une base est indiquée plus bas.

C. k-formes.

Définition. On appelle *forme extérieure de degré k* ou *k-forme* une fonction k-linéaire et antisymétrique de k vecteurs:

$$\omega(\lambda_1\xi_1' + \lambda_2\xi_1'', \xi_2, \ldots, \xi_k) = \lambda_1\omega(\xi_1', \xi_2, \ldots, \xi_k) + \lambda_2\omega(\xi_1'', \xi_2, \ldots, \xi_k),$$

$$\omega(\xi_{i_1}, \ldots, \xi_{i_k}) = (-1)^\nu \omega(\xi_1, \ldots, \xi_k), \text{ où } \nu = \begin{cases} 0 \text{ si la permutation } i_1, \ldots, i_k \text{ est paire;} \\ 1 \text{ si elle est impaire.} \end{cases}$$

Exemple 1. Le volume orienté d'un parallélépipède d'arêtes $\xi_1, \ldots, \xi_n$ dans l'espace euclidien orienté R^n est une n-forme (fig. 138)

$$V(\xi_1, \ldots, \xi_n) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{1n} \\ \vdots & & \\ \xi_{n1} & \cdots & \xi_{nn} \end{vmatrix},$$

où $\xi_i = \xi_{i1}e_1 + \ldots + \xi_{in}e_n$ et $e_1, \ldots, e_n$ est une base dans R^n.

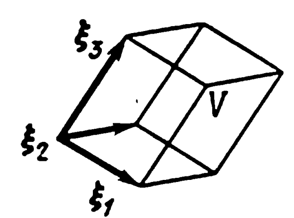


Fig. 138. Le volume orienté est une 3-forme.

Exemple 2. Soit R^k un plan orienté de dimension k dans l'espace euclidien R^n. Alors le volume orienté de dimension k de la projection d'un parallélépipède d'arêtes $\xi_1, \xi_2, \ldots, \xi_k \in R^n$ sur R^k est une k-forme sur R^n.

L'ensemble de toutes les k-formes sur $\mathbf{R}^n$ se transforme en espace vectoriel réel si l'on y définit la somme comme

$$(\omega_1 + \omega_2)(\xi) = \omega_1(\xi) + \omega_2(\xi),$$
$$\xi = \{\xi_1, \ldots, \xi_k\}, \quad \xi_i \in \mathbf{R}^n,$$

et le produit d'une k-forme par un nombre comme

$$(\lambda\omega)(\xi) = \lambda\omega(\xi).$$

Exercice 3. Montrer que cet espace vectoriel est de dimension finie et trouver cette dimension.

Réponse. C_n^k; une base est indiquée plus bas.

D. Produit extérieur de deux 1-formes. Introduisons encore une opération : le produit extérieur des formes. Si ω^k est une k-forme et ω^l une l-forme sur $\mathbf{R}^n$, alors leur produit extérieur $\omega^k \wedge \omega^l$ est une $(k+l)$-forme. Définissons d'abord le produit extérieur de 1-formes en associant à chaque couple de 1-formes ω_1, ω_2 sur $\mathbf{R}^n$ une 2-forme $\omega_1 \wedge \omega_2$ sur $\mathbf{R}^n$.

Soit ξ un vecteur de $\mathbf{R}^n$. Si l'on a deux 1-formes ω_1 et ω_2, on peut définir une application de $\mathbf{R}^n$ dans le plan $\mathbf{R} \times \mathbf{R}$ en associant à $\xi \in \mathbf{R}^n$ le vecteur $\omega(\xi)$ de composantes $\omega_1(\xi)$, $\omega_2(\xi)$ sur le plan décrit par les coordonnées ω_1, ω_2 (fig. 139).

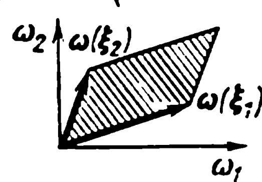


Fig. 139. Détermination du produit extérieur de deux 1-formes.

Définition. La *valeur du produit extérieur* $\omega_1 \wedge \omega_2$ *sur un couple de vecteurs* $\xi_1, \xi_2 \in \mathbf{R}^n$ est égale à l'aire orientée de l'image du parallélogramme construit sur les vecteurs ξ_1, ξ_2 dans le plan (ω_1, ω_2) :

$$(\omega_1 \wedge \omega_2)(\xi_1, \xi_2) = \begin{vmatrix} \omega_1(\xi_1) & \omega_2(\xi_1) \\ \omega_1(\xi_2) & \omega_2(\xi_2) \end{vmatrix}.$$

Exercice 4. Montrer que $\omega_1 \wedge \omega_2$ est bien une 2-forme.

Exercice 5. Montrer que l'application

$$(\omega_1, \omega_2) \to \omega_1 \wedge \omega_2$$

est bilinéaire et antisymétrique :

$$\omega_1 \wedge \omega_2 = -\omega_2 \wedge \omega_1,$$
$$(\lambda'\omega_1' + \lambda''\omega_1'') \wedge \omega_2 = \lambda'\omega_1' \wedge \omega_2 + \lambda''\omega_1'' \wedge \omega_2.$$

Indication. Le déterminant est bilinéaire et antisymétrique suivant les lignes et les colonnes.

Supposons donné sur $\mathbf{R}^n$ un système de coordonnées linéaires, i.e. n 1-formes indépendantes $x_1, \ldots, x_n$ que nous appellerons *formes bases*.

Les produits extérieurs des bases sont des 2-formes $x_i \wedge x_j$. En vertu de l'antisymétrie $x_i \wedge x_i = 0$, $x_i \wedge x_j = -x_j \wedge x_i$. La forme $x_i \wedge x_j$ admet une interprétation géométrique simple: elle est égale sur le couple de vecteurs ξ_1, ξ_2 à l'aire orientée de la projection du parallélogramme ξ_1, ξ_2 sur le plan décrit par x_i, x_j parallèlement aux directions des autres coordonnées.

Exercice 6. Montrer que les $C_n^2 = \frac{n(n-1)}{n}$ formes $x_i \wedge x_j$ $(i < j)$ sont linéairement indépendantes.

En particulier, dans l'espace engendré par (x_1, x_2, x_3), l'aire de la projection sur le plan (x_1, x_2) est le produit $x_1 \wedge x_2$, sur le plan (x_2, x_3), le produit $x_2 \wedge x_3$ et sur le plan (x_3, x_1), le produit $x_3 \wedge x_1$.

Exercice 7. Montrer que

$$Px_2 \wedge x_3 + Qx_3 \wedge x_1 + Rx_1 \wedge x_2$$

est l'expression de toutes les 2-formes de l'espace (x_1, x_2, x_3).

Exercice 8. Montrer que toute 2-forme de l'espace de dimension n muni des coordonnées $x_1, \ldots, x_n$ s'écrit de façon unique

$$\omega^2 = \sum_{i<j} a_{ij} x_i \wedge x_j.$$

Indication. Soit e_i le i-ème vecteur de base, i.e. $x_i(e_i) = 1$, $x_j(e_i) = 0$, $j \neq i$. Considérons les valeurs de la forme ω^2 sur le couple e_i, e_j. On a

$$a_{ij} = \omega^2(e_i, e_j).$$

E. Monômes extérieurs. Soient données k 1-formes $\omega_1, \ldots, \omega_k$. Définissons leur produit extérieur $\omega_1 \wedge \ldots \wedge \omega_k$.

Définition. Posons

$$(\omega_1 \wedge \ldots \wedge \omega_k)(\xi_1, \ldots, \xi_k) = \begin{vmatrix} \omega_1(\xi_1) & \ldots & \omega_k(\xi_1) \\ \omega_1(\xi_k) & \ldots & \omega_k(\xi_k) \end{vmatrix}.$$

Autrement dit, *la valeur du produit de 1-formes sur le parallélépipède $\xi_1, \ldots, \xi_k$ est égale au volume orienté de l'image de ce parallélépipède dans l'espace arithmétique euclidien orienté $\mathbf{R}^k$ par l'application $\xi \to (\omega_1(\xi), \ldots, \omega_k(\xi))$.*

Exercice 9. Montrer que $\omega_1 \wedge \ldots \wedge \omega_k$ est une k-forme.

Exercice 10. Montrer que le produit extérieur de 1-formes définit une application multilinéaire antisymétrique

$$(\omega_1, \ldots, \omega_k) \to \omega_1 \wedge \ldots \wedge \omega_k.$$

Autrement dit,

$$(\lambda'\omega_1' + \lambda''\omega_1'') \wedge \omega_2 \wedge \ldots \wedge \omega_k = \\ = \lambda'\omega_1' \wedge \omega_2 \wedge \ldots \wedge \omega_k + \lambda''\omega_1'' \wedge \omega_2 \wedge \ldots \wedge \omega_k,$$

$$\omega_{i_1} \wedge \ldots \wedge \omega_{i_k} = (-1)^\nu \omega_1 \wedge \ldots \wedge \omega_k, \text{ où } \nu = \begin{cases} 0 \text{ si la permutation } i_1, \ldots, i_k \text{ est paire,} \\ 1 \text{ si elle est impaire.} \end{cases}$$

Considérons maintenant sur $\mathbf{R}^n$ un système de coordonnées défini par les formes bases $x_1, \ldots, x_n$. Le produit extérieur de k formes bases

$$x_{i_1} \wedge \ldots \wedge x_{i_k}, \quad 1 \leqslant i_m \leqslant n,$$

est le volume orienté de la projection d'un parallélépipède de dimension k sur le plan $(x_{i_1}, \ldots, x_{i_k})$ de dimension k parallèlement aux directions des autres coordonnées.

Exercice 11. Montrer que si parmi les indices $i_1, \ldots, i_k$ il en est deux qui sont identiques, la forme $x_{i_1} \wedge \ldots \wedge x_{i_k}$ est nulle.

Exercice 12. Montrer que les formes

$$x_{i_1} \wedge \ldots \wedge x_{i_k}, \text{ où } 1 \leqslant i_1 < \ldots < i_k \leqslant n,$$

sont linéairement indépendantes.

Le nombre de ces formes est visiblement C_n^k. Nous les appellerons *k-formes bases.*

Exercice 13. Montrer que toute k-forme sur $\mathbf{R}^n$ est représentée de façon unique par une combinaison linéaire de k-formes bases:

$$\omega^k = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} a_{i_1 \ldots i_k} x_{i_1} \wedge \ldots \wedge x_{i_k}.$$

Indication. $a_{i_1 \ldots i_k} = \omega^k (e_{i_1}, \ldots, e_{i_k})$.

Il résulte de cet exercice que l'espace vectoriel des k-formes sur $\mathbf{R}^n$ est de dimension C_n^k. En particulier, pour $k = n$ on a $C_n^k = 1$, d'où il vient le

Corollaire. *Toute n-forme sur* $\mathbf{R}^n$ *soit est le volume orienté d'un parallélépipède pour une certaine unité de volume, soit est nulle*:

$$\omega^n = a \cdot x_1 \wedge \ldots \wedge x_n.$$

Exercice 14. Montrer que toute k-forme sur $\mathbf{R}^n$ est nulle pour $k > n$.

Passons maintenant au produit d'une k-forme ω^k par une l-forme ω^l. Soient donnés tout d'abord deux monômes

$$\omega^k = \omega_1 \wedge \ldots \wedge \omega_k, \quad \omega^l = \omega_{k+1} \wedge \ldots \wedge \omega_{k+l},$$

où $\omega_1, \ldots, \omega_{k+l}$ sont des 1-formes. Définissons leur produit $\omega^k \wedge \omega^l$ comme le monôme

$$(\omega_1 \wedge \ldots \wedge \omega_k) \wedge (\omega_{k+1} \wedge \ldots \wedge \omega_{k+l}) = \\ = \omega_1 \wedge \ldots \wedge \omega_k \wedge \omega_{k+1} \wedge \ldots \wedge \omega_{k+l}.$$

Exercice 15. Montrer que le produit des monômes est associatif

$$(\omega^k \wedge \omega^l) \wedge \omega^m = \omega^k \wedge (\omega^l \wedge \omega^m)$$

et anticommutatif:

$$\omega^k \wedge \omega^l = (-1)^{kl} \omega^l \wedge \omega^k.$$

I n d i c a t i o n. Pour permuter chacun des l facteurs ω^l à la première place il faut effectuer k inversions avec les k facteurs ω^k.

R e m a r q u e. Il est utile de se rappeler que le produit anticommutatif est symétrique si l'un au moins des exposants k et l est pair, et antisymétrique si les deux exposants k et l sont impairs.

§ 33. Produit extérieur

On définit ici le produit extérieur de formes et on montre qu'il est anticommutatif, distributif et associatif.

A. Définition du produit extérieur. Nous allons maintenant définir le produit extérieur d'une k-forme arbitraire ω^k par une l-forme ω^l. Le produit $\omega^k \wedge \omega^l$ est une $(k+l)$-forme. On verra que le produit est

1) anticommutatif: $\omega^k \wedge \omega^l = (-1)^{kl}\omega^l \wedge \omega^k$,

2) distributif: $(\lambda_1\omega_1^k + \lambda_2\omega_2^k) \wedge \omega^l = \lambda_1\omega_1^k \wedge \omega^l + \lambda_2\omega_2^k \wedge \omega^l$,

3) associatif: $(\omega^k \wedge \omega^l) \wedge \omega^m = \omega^k \wedge (\omega^l \wedge \omega^m)$.

D é f i n i t i o n. Le *produit extérieur* $\omega^k \wedge \omega^l$ *d'une k-forme* ω^k *sur* $\mathbf{R}^n$ *par une l-forme* ω^l *sur* $\mathbf{R}^n$ est la $(k+l)$-forme sur $\mathbf{R}^n$ dont la valeur sur les $k+l$ vecteurs $\xi_1, \ldots, \xi_k, \xi_{k+1}, \ldots, \xi_{k+l} \in \mathbf{R}^n$ est égale à

$$(\omega^k \wedge \omega^l)(\xi_1, \ldots, \xi_{k+l}) = \sum (-1)^\nu \omega^k \times \times (\xi_{i_1}, \ldots, \xi_{i_k})\, \omega^l (\xi_{j_1}, \ldots, \xi_{j_l}), \quad (1)$$

où $i_1 < \ldots < i_k$, $j_1 < \ldots < j_l$; $(i_1, \ldots, i_k, j_1, \ldots, j_l)$ est une permutation des indices $(1, 2, \ldots, k+l)$ et

$$\nu = \begin{cases} 1 \text{ si cette permutation est impaire,} \\ 0 \text{ si elle est paire.} \end{cases}$$

Autrement dit, *toute partition des $k+l$ vecteurs $\xi_1, \ldots, \xi_{k+l}$ en deux groupes (l'un de k vecteurs et l'autre de l) définit un terme de la somme* (1). *Ce terme est égal au produit de la valeur prise par ω^k sur les k vecteurs du premier groupe par la valeur prise par ω^l sur les l vecteurs du second groupe avec le signe + ou — selon l'ordre des vecteurs dans les groupes*: *le signe + si la permutation $\xi_1, \xi_2, \ldots, \xi_{k+l}$ des k vecteurs du premier groupe et des l vecteurs du second écrits les uns à la suite des autres est paire et le signe — si elle est impaire.*

E x e m p l e. Si $k = 2$, on a deux partitions en tout: ξ_1, ξ_2 et ξ_2, ξ_1. Donc

$$(\omega_1 \wedge \omega_2)(\xi_1, \xi_2) = \omega_1(\xi_1)\,\omega_2(\xi_2) - \omega_2(\xi_1)\,\omega_1(\xi_2),$$

en vertu de la définition du produit de 1-formes (§ 32).

E x e r c i c e 1. Montrer que la définition donnée plus haut définit réellement une $(k+l)$-forme (i.e. la valeur $(\omega^k \wedge \omega^l)(\xi_1, \ldots, \xi_{k+l})$ est multilinéaire et antisymétrique en ξ).

B. Propriétés du produit extérieur.

Théorème. *Le produit extérieur défini plus haut est anticommutatif, distributif et associatif. Pour les monômes il coïncide avec le produit défini dans le* § 32.

La démonstration de l'anticommutativité est basée sur les propriétés élémentaires des permutations paires et impaires (cf. exercice de la fin du § 32). On la laisse au soin du lecteur.

La distributivité découle du fait que tout terme de (1) est linéaire en ω^k et ω^l.

La démonstration de l'associativité fait appel à l'analyse combinatoire; nous allons nous servir du théorème de Laplace sur le développement d'un déterminant suivant les mineurs des colonnes *).

Commençons par la remarque suivante: *Si l'associativité est démontrée pour les termes d'une somme, elle est valable pour cette somme*, i. e.

$$\left.\begin{array}{l}(\omega_1' \wedge \omega_2) \wedge \omega_3 = \omega_1' \wedge (\omega_2 \wedge \omega_3) \\ (\omega_1'' \wedge \omega_2) \wedge \omega_3 = \omega_1'' \wedge (\omega_2 \wedge \omega_3)\end{array}\right\} \Rightarrow ((\omega_1' + \omega_1'') \wedge \omega_2) \wedge \omega_3 =$$
$$= (\omega_1' + \omega_1'') \wedge (\omega_2 \wedge \omega_3).$$

Car en vertu de la distributivité qui a déjà été démontrée on a

$$((\omega_1' + \omega_1'') \wedge \omega_2) \wedge \omega_3 = ((\omega_1' \wedge \omega_2) \wedge \omega_3) + ((\omega_1'' \wedge \omega_2) \wedge \omega_3),$$
$$(\omega_1' + \omega_1'') \wedge (\omega_2 \wedge \omega_3) = (\omega_1' \wedge (\omega_2 \wedge \omega_3)) + (\omega_1'' \wedge (\omega_2 \wedge \omega_3)).$$

Or nous savons déjà (§ 32, exercice 12) que toute forme sur $\mathbf{R}^n$ est une somme de monômes. Donc il suffit de démontrer l'associativité du produit de monômes.

Comme nous n'avons pas encore démontré l'équivalence de la définition du produit de k 1-formes du § 32 avec la définition générale (1), nous désignerons provisoirement le produit de k 1-formes par le symbole $\overline{\wedge}$, de sorte que nos monômes s'écrivent

$$\omega^k = \omega_1 \overline{\wedge} \ldots \overline{\wedge} \omega_k, \quad \omega^l = \omega_{k+1} \overline{\wedge} \ldots \overline{\wedge} \omega_{k+l},$$

où $\omega_1, \ldots, \omega_{k+l}$ sont des 1-formes.

---

*) La démonstration immédiate de l'associativité (qui généralise également celle du théorème de Laplace) consiste à vérifier la concordance des signes dans l'identité

$$((\omega^k \wedge \omega^l) \wedge \omega^m)(\xi_1, \ldots, \xi_{k+l+m}) =$$
$$= \sum \pm \omega^k (\xi_{i_1}, \ldots, \xi_{i_k}) \omega^l (\xi_{j_1}, \ldots, \xi_{j_l}) \omega^m (\xi_{h_1}, \ldots, \xi_{h_m}),$$

où $i_1 < \ldots < i_k$, $j_1 < \ldots < j_l$, $h_1 < \ldots < h_m$; $(i_1, \ldots, h_m)$ est une permutation des nombres $(1, \ldots, k+l+m)$.

Lemme. *Le produit extérieur de deux monômes est un monôme :*

$$(\omega_1 \overline{\wedge} \dots \overline{\wedge} \omega_k) \wedge (\omega_{k+1} \overline{\wedge} \dots \overline{\wedge} \omega_{k+l}) =$$
$$= \omega_1 \overline{\wedge} \dots \overline{\wedge} \omega_k \overline{\wedge} \omega_{k+1} \overline{\wedge} \dots \overline{\wedge} \omega_{k+l}.$$

Démonstration. Calculons les valeurs prises par le premier et le second membre sur les $k + l$ vecteurs $\xi_1, \dots, \xi_{k+l}$. Le second membre est égal au déterminant $\det |\omega_i(\xi_j)|$ d'ordre $k + l$. En vertu de la formule (1) le premier membre est égal à la somme des produits

$$\sum \pm \det_{1 \leqslant i \leqslant k} |\omega_i(\xi_{i_m})| \cdot \det_{k < i \leqslant k+l} |\omega_i(\xi_{j_m})|$$

des mineurs des k premières colonnes du déterminant d'ordre $k + l$ par les mineurs complémentaires. Le théorème de Laplace sur le développement suivant les mineurs relatifs aux k premières colonnes affirme justement que cette somme, dans laquelle les signes sont choisis comme dans la définition (1), est égale à $\det |\omega_i(\xi_j)|$. Le lemme est démontré.

Ce lemme entraîne l'équivalence des opérations $\overline{\wedge}$ et $\wedge$. En effet, nous obtenons successivement

$$\omega_1 \overline{\wedge} \omega_2 = \omega_1 \wedge \omega_2,$$

$$\omega_1 \overline{\wedge} \omega_2 \overline{\wedge} \omega_3 = (\omega_1 \overline{\wedge} \omega_2) \wedge \omega_3 = (\omega_1 \wedge \omega_2) \wedge \omega_3,$$

$$\omega_1 \overline{\wedge} \omega_2 \overline{\wedge} \dots \overline{\wedge} \omega_k = (\dots((\omega_1 \wedge \omega_2) \wedge \omega_3) \wedge \dots \wedge \omega_k).$$

L'associativité évidente du produit $\overline{\wedge}$ de k 1-formes entraîne l'associativité du produit $\wedge$ des monômes. Ce qui, en vertu de la remarque faite plus haut, démontre l'associativité dans le cas général.

Exercice 2. Montrer que le carré extérieur d'une 1-forme ou, d'une façon générale, d'une forme de degré impair est nul : $\omega^k \wedge \omega^k = 0$ si k est impair.

Exemple 1. Soit un système de coordonnées $p_1, \dots, p_n, q_1, \dots, q_n$ et une 2-forme $\omega^2 = \sum_{i=1}^{n} p_i \wedge q_i$ sur R^{2n}.

(Géométriquement la forme ω^2 désigne la somme des aires orientées des projections d'un parallélogramme sur n plans arithmétiques de dimension deux $(p_1, q_1); \dots; (p_n, q_n)$. Dans la suite nous verrons que la 2-forme ω^2 est fondamentale en mécanique hamiltonienne. On peut démontrer que toute 2-forme régulière *) sur R^{2n} s'écrit comme ω^2 dans un certain système de coordonnées $(p_1, \dots, q_n)$.)

Exercice 3. Trouver le carré extérieur de la 2-forme ω^2.

Réponse. $\omega^2 \wedge \omega^2 = -2 \sum_{i>j} p_i \wedge p_j \wedge q_i \wedge q_j$.

---

*) La forme bilinéaire ω^2 est régulière si $\forall \xi \neq 0, \exists \eta : \omega^2(\xi, \eta) \neq 0$.

Exercice 4. Trouver la puissance extérieure k de ω^2.

Réponse. $\underbrace{\omega^2 \wedge \omega^2 \wedge \ldots \wedge \omega^2}_{k} = \pm k! \sum_{i_1 < \ldots < i_k} p_{i_1} \wedge \ldots \wedge p_{i_k} \wedge q_{i_1} \wedge \ldots$ $\ldots \wedge q_{i_k}$. En particulier

$$\underbrace{\omega^2 \wedge \ldots \wedge \omega^2}_{n} = \pm n! p_1 \wedge \ldots \wedge p_n \wedge q_1 \wedge \ldots \wedge q_n$$

est au facteur $\pm n!$ près le volume d'un $2n$-parallélépipède de R^{2n}.

Exemple 2. Considérons maintenant l'espace euclidien orienté R^3. A tout vecteur $A \in R^3$ associons la 1-forme ω_A^1 en posant

$$\omega_A^1(\xi) = (A, \xi) \text{ (produit scalaire)}$$

et la 2-forme ω_A^2 en posant

$$\omega_A^2(\xi_1, \xi_2) = (A, \xi_1, \xi_2) \text{ (produit mixte).}$$

Exercice 5. Montrer que les applications $A \to \omega_A^1$ et $A \to \omega_A^2$ sont des isomorphismes de l'espace vectoriel R^3 des vecteurs A sur les espaces vectoriels des 1-formes sur R^3 et des 2-formes sur R^3. Si dans R^3 on fait choix d'une base orthonormée orientée (x_1, x_2, x_3), alors

$$\omega_A^1 = A_1x_1 + A_2x_2 + A_3x_3, \qquad \omega_A^2 = A_1x_2 \wedge x_3 + A_2x_3 \wedge x_1 + A_3x_1 \wedge x_2.$$

Remarque. Ainsi les isomorphismes indiqués ne dépendent pas du choix de la base orthonormée orientée (x_1, x_2, x_3). Cependant ils dépendent du choix de la structure euclidienne sur R^3, quant à l'isomorphisme $A \to \omega_A^2$ il dépend également de l'orientation (qui figure implicitement dans la définition du produit mixte).

Exercice 6. Montrer qu'en vertu des isomorphismes établis le produit extérieur de 1-formes se transforme en produit vectoriel de vecteurs de R^3, i.e.

$$\omega_A^1 \wedge \omega_B^1 = \omega_{[A, B]}^2 \text{ quels que soient } A, B \in R^3.$$

Donc le *produit extérieur de formes peut être considéré comme une généralisation du produit vectoriel sur* R^3 *au cas multidimensionnel.* Cependant dans le cas multidimensionnel le produit n'est pas un vecteur du même espace: l'espace des 2-formes sur R^n n'est isomorphe à R^n que pour $n = 3$.

Exercice 7. Montrer qu'en vertu des isomorphismes établis plus haut le produit extérieur d'une 1-forme par une 2-forme se transforme en un produit scalaire de vecteurs de R^3:

$$\omega_A^1 \wedge \omega_B^2 = (A, B)\, x_1 \wedge x_2 \wedge x_3.$$

C. Comportement par les applications. Soient $f: R^m \to R^n$ une application linéaire, ω^k une k-forme extérieure sur R^n. Alors sur R^m est définie une k-forme $f^*\omega^k$ qui prend sur k vecteurs $\xi_1, \ldots, \xi_k \in R^n$ une valeur égale à celle que prend ω^k sur leurs images:

$$(f^*\omega^k)(\xi_1, \ldots, \xi_k) = \omega^k(f\xi_1, \ldots, f\xi_k).$$

Exercice 8. Vérifier que $f^*\omega^k$ est une forme extérieure.

Exercice 9. Vérifier que f^* est un opérateur linéaire de l'espace des

k-formes sur $\mathbf{R}^n$ dans l'espace des k-formes sur $\mathbf{R}^m$ (l'astérisque indique que f^* opère dans le sens *inverse* de f).

Exercice 10. Soit $f: \mathbf{R}^m \to \mathbf{R}^n$, $g: \mathbf{R}^n \to \mathbf{R}^p$. Vérifier que $(g \circ f)^* = f^* \circ g^*$.

Exercice 11. Vérifier que f^* conserve le produit extérieur:

$$f^* (\omega^k \wedge \omega^l) = (f^* \omega^k) \wedge (f^* \omega^l).$$

§ 34. Formes différentielles

On donne ici la définition d'une forme différentielle sur une variété différentiable.

A. 1-formes différentielles. Un exemple simple de forme différentielle est la différentielle d'une fonction.

Exemple. Considérons la fonction $y = f(x) = x^2$. Sa différentielle $df = 2x\, dx$ dépend du point x et de l'« accroissement de l'argument », i.e. du vecteur ξ tangent à l'axe des x. Fixons le point x. La différentielle de la fonction

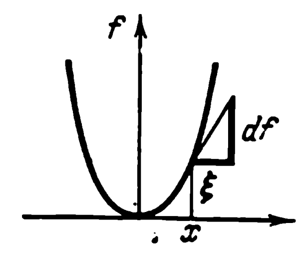


Fig. 140. Différentielle d'une fonction.

Fig. 141. Problème 1.

au point x, $df|_x$, dépend *linéairement* de ξ. Si donc $x = 1$ et la coordonnée du vecteur tangent ξ est égale à 1, alors $df = 2$, et si la coordonnée de ξ est égale à 10, alors $df = 20$ (fig. 140).

Soit $f: M \to \mathbf{R}$ une fonction différentiable définie sur une variété M (on peut s'imaginer une « fonction de plusieurs variables » $f: \mathbf{R}^n \to \mathbf{R}$). La différentielle $df|_x$ de la fonction f en un point x est une application linéaire

$$df_x: TM_x \to \mathbf{R}$$

de l'espace tangent à M au point x dans la droite réelle. Rappelons la définition de cette application.

Soit $\xi \in TM_x$ le vecteur vitesse de la courbe $x(t): \mathbf{R} \to M$, $x(0) = x$, $\dot{x}(0) = \xi$. Alors par définition:

$$df_x(\xi) = \frac{d}{dt}\Big|_{t=0} f(x(t)).$$

Exercice 1. Soit ξ le vecteur vitesse de la courbe plane $x(t) = \cos t$, $y(t) = \sin t$ pour $t = 0$. Calculer les différentielles dx et dy des fonctions x et y sur le vecteur ξ (fig. 141).

Réponse.

$$dx_1|_{1,0}(\xi) = 0, \quad dy|_{1,0}(\xi) = 1.$$

Remarquons que la différentielle de la fonction f en un point $x \in M$ est la 1-forme df_x sur l'espace tangent TM_x.

La différentielle df de la fonction f sur la variété M est une application du fibré tangent TM dans la droite $\mathbf{R}$

$$df: TM \to \mathbf{R} \ (TM = \bigcup_x TM_x).$$

Cette application est linéaire sur tout espace tangent $TM_x \subset TM$ et différentiable.

Définition. *On appelle forme différentielle de degré* 1 (*ou* 1-*forme*) *sur une variété M* une application différentiable

$$\omega : TM \to \mathbf{R}$$

du fibré tangent TM à la variété M dans la droite $\mathbf{R}$, linéaire sur tout espace tangent TM_x.

On peut dire qu'une 1-*forme différentielle sur M est une* 1-*forme algébrique sur TM_x* « différentiable par rapport à x ».

Exercice 2. Montrer que toute 1-forme différentielle sur une droite est différentielle d'une certaine fonction.

Exercice 3. Trouver sur le cercle et le plan des 1-formes différentielles qui ne soient différentielles d'aucune fonction.

B. Forme générale des 1-formes différentielles sur $\mathbf{R}^n$. Prenons pour variété M un espace vectoriel muni des coordonnées $x_1, \ldots, x_n$. Rappelons que les composantes $\xi_1, \ldots, \xi_n$ d'un vecteur tangent $\xi \in T\mathbf{R}^n_x$ sont les valeurs prises par les différentielles des coordonnées $dx_1, \ldots, dx_n$ sur le vecteur ξ. Ces n 1-formes sur $T\mathbf{R}^n_x$ sont linéairement indépendantes. Donc les 1-formes $dx_1, \ldots, dx_n$ forment une base dans l'espace n-dimensionnel des 1-formes sur $T\mathbf{R}^n_x$.

Par conséquent toute 1-forme sur $T\mathbf{R}^n_x$ s'écrit de manière unique $a_1\, dx_1 + \ldots + a_n\, dx_n$, où a_i sont des coefficients réels. Soit maintenant ω une 1-forme différentielle quelconque sur $\mathbf{R}^n$. En chaque point x elle se décompose de façon unique suivant la base $dx_1, \ldots, dx_n$. D'où il s'ensuit le

Fig. 142. Problème 4.

Théorème. *Toute 1-forme différentielle sur $\mathbf{R}^n$ dans lequel on a choisi un système de coordonnées $x_1, \ldots, x_n$ s'écrit de façon unique*

$$\omega = a_1(x)\,dx_1 + \ldots + a_n(x)\,dx_n,$$

où les coefficients $a_i(x)$ sont des fonctions différentiables.

E x e r c i c e 4. Calculer les formes $\omega_1 = dx_1$, $\omega_2 = x_1\,dx_2$, $\omega_3 = dr^2$ ($r^2 = x_1^2 + x_2^2$) sur les vecteurs ξ_1, ξ_2, ξ_3 (fig. 142).

Réponse.

	ξ_1	ξ_2	ξ_3
ω_1	0	−1	1
ω_2	0	−2	−2
ω_3	0	−8	0

E x e r c i c e 5. Soient $x_1, \ldots, x_n$ des fonctions formant un système de coordonnées local dans un domaine d'une variété M. Montrer que toute 1-forme sur ce domaine s'écrit de façon unique $\omega = a_1(x)\,dx_1 + \ldots + a_n(x)dx_n$.

C. k-formes différentielles.

D é f i n i t i o n. *On appelle k-forme différentielle $\omega^k|_x$ au point x d'une variété M une k-forme extérieure sur l'espace TM_x tangent à M en x, i. e. une fonction k-linéaire antisymétrique de k vecteurs $\xi_1, \ldots, \xi_k$ tangents à M en x.*

Si une telle forme $\omega^k|_x$ est donnée en chaque point x de M et si elle est différentiable, on dit qu'est donnée une *k-forme ω^k sur la variété M.*

E x e r c i c e 6. Munir d'une structure naturelle de variété différentiable un ensemble dont chaque élément est une collection de k vecteurs tangents à M en un point quelconque x.

Une k-forme différentielle est une application différentiable de la variété obtenue dans une droite.

On peut dire qu'une *k-forme sur M est une k-forme extérieure sur TM_x « dépendant différentiablement de x ».*

L'addition, la multiplication par un nombre et le produit extérieur de formes sur M se définissent ponctuellement: en chaque point $x \in M$ il faut additionner, multiplier par un nombre ou faire le produit extérieur des formes algébriques extérieures sur TM_x.

E x e r c i c e 7. Montrer que toutes les k-formes sur M forment un espace vectoriel (de dimension infinie si k n'est pas supérieur à la dimension de M).

Les formes différentielles peuvent être multipliées par des nombres et par des fonctions. L'ensemble C^∞ des k-formes différentielles possède donc une structure naturelle de module sur l'anneau des fonctions réelles indéfiniment différentiables sur M.

E. Forme générale des k-formes différentielles sur $\mathbf{R}^n$. Prenons pour variété M l'espace vectoriel $\mathbf{R}^n$ dans lequel on aura fixé des coordonnées $x_1, \ldots, x_n : \mathbf{R}^n \to \mathbf{R}$. Fixons un point x. Nous avons vu plus haut que les n 1-formes $dx_1, \ldots, dx_n$ forment une base dans l'espace des 1-formes sur l'espace tangent $T\mathbf{R}^n_x$.

Considérons les produits extérieurs des formes bases

$$dx_{i_1} \wedge \ldots \wedge dx_{i_k}, \quad i_1 < \ldots < i_k.$$

Au § 32 nous avons vu que ces C_n^k k-formes engendraient une base dans l'espace des k-formes extérieures sur $T\mathbf{R}^n_x$. Donc toute k-forme extérieure sur $T\mathbf{R}^n_x$ s'écrit de façon unique

$$\sum_{i_1 < \ldots < i_k} a_{i_1 \ldots i_k} dx_{i_1} \wedge \ldots \wedge dx_{i_k}.$$

Soit maintenant ω^k une k-forme différentielle arbitraire sur $\mathbf{R}^n$. En chaque point x elle se décompose de façon unique suivant la base mentionnée plus haut. Il résulte donc le

Théorème. *Toute k-forme différentielle sur $\mathbf{R}^n$ dans lequel on a choisi un système de coordonnées $x_1, \ldots, x_n$ s'écrit de façon unique*

$$\omega^k = \sum_{i_1 < \ldots < i_k} a_{i_1 \ldots i_k}(x)\, dx_{i_1} \wedge \ldots dx_{i_k},$$

où $a_{i_1 \ldots i_k}(x)$ sont des fonctions différentiables sur $\mathbf{R}^n$.

E x e r c i c e 8. Calculer les formes $\omega_1 = dx_1 \wedge dx_2$, $\omega_2 = x_1 dx_1 \wedge dx_2 -$

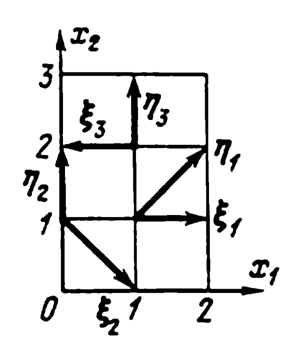


Fig. 143. Problème 8.

$- x_2 dx_2 \wedge dx_1$, $\omega_3 = r\, dr \wedge d\varphi$ (où $x_1 = r \cos \varphi$, $x_2 = r \sin \varphi$) sur les couples de vecteurs (ξ_1, η_1), (ξ_2, η_2), (ξ_3, η_3) (fig. 143).

Réponse.

	(ξ_1, η_1)	(ξ_2, η_2)	(ξ_3, η_3)
ω_1	1	1	−1
ω_2	2	1	−3
ω_3	1	1	−1

Exercice 9. Calculer les formes $\omega_1 = dx_2 \wedge dx_3$, $\omega_2 = x_1\, dx_3 \wedge dx_2$, $\omega_3 = dx_3 \wedge dr^2$, $r^2 = x_1^2 + x_2^2 + x_3^2$, sur le couple de vecteurs $\xi = (1, 1, 1)$, $\eta = (1, 2, 3)$ appliqués au point $x = (2, 0, 0)$.

Réponse. $\omega_1 = 1$, $\omega_2 = -2$, $\omega_3 = -8$.

Exercice 10. Soient $x_1, \ldots, x_n : M \to \mathbf{R}$ des fonctions sur M formant un système de coordonnées local dans un certain domaine. Montrer que toute forme différentielle sur ce domaine s'écrit de façon unique

$$\omega^k = \sum_{i_1 < \ldots < i_k} a_{i_1 \ldots i_k}(x)\, dx_{i_1} \wedge \ldots \wedge dx_{i_k}.$$

Exemple. *Changement de variables dans une forme.* Soient donnés deux systèmes de coordonnées x_1, x_2, x_3 et y_1, y_2, y_3 et une 2-forme ω sur $\mathbf{R}^3$. En vertu du dernier théorème, dans les coordonnées x, la forme ω s'écrit $\omega = X_1\, dx_2 \wedge dx_3 + X_2\, dx_3 \wedge dx_1 + X_3\, dx_1 \wedge dx_2$, où X_1, X_2, X_3 sont des fonctions de x_1, x_2, x_3 et dans les coordonnées y, sous la forme $\omega = Y_1\, dy_2 \wedge dy_3 + Y_2\, dy_3 \wedge dy_1 + Y_3\, dy_1 \wedge dy_2$, où Y_1, Y_2, Y_3 sont des fonctions de y_1, y_2, y_3.

Exercice 11. Connaissant l'écriture d'une forme dans les coordonnées x (i.e. les X_i) et la formule de changement de variables $x = x(y)$ écrire cette forme dans les coordonnées y, i.e. trouver les fonctions Y_i.

Solution. On a $dx_i = \frac{\partial x_i}{\partial y_1} dy_1 + \frac{\partial x_i}{\partial y_2} dy_2 + \frac{\partial x_i}{\partial y_3} dy_3$. Donc

$$dx_2 \wedge dx_3 = \left(\frac{\partial x_2}{\partial y_1} dy_1 + \frac{\partial x_2}{\partial y_2} dy_2 + \frac{\partial x_2}{\partial y_3} dy_3 \right) \wedge$$
$$\wedge \left(\frac{\partial x_3}{\partial y_1} dy_1 + \frac{\partial x_3}{\partial y_2} dy_2 + \frac{\partial x_3}{\partial y_3} dy_3 \right),$$

d'où

$$Y_3 = X_1 \left| \frac{D(x_2, x_3)}{D(y_1, y_2)} \right| + X_2 \left| \frac{D(x_3, x_1)}{D(y_1, y_2)} \right| + X_3 \left| \frac{D(x_1, x_2)}{D(y_1, y_2)} \right|, \text{ etc.}$$

F. Appendice. Formes différentielles sur un espace de dimension trois. Soit M une variété riemannienne orientée de dimension trois (dans tous les exemples qui vont suivre M est assimilé à l'espace euclidien ordinaire $\mathbf{R}^3$). Soient x_1, x_2, x_3 des coordonnées locales et supposons que le carré de l'élément de longueur soit de la forme

$$ds^2 = E_1\, dx_1^2 + E_2\, dx_2^2 + E_3\, dx_3^2$$

(i. e. le système de coordonnées est orthogonal).

Exercice 12. Calculer E_1, E_2, E_3 pour les coordonnées cartésiennes x, y, z, les coordonnées cylindriques r, φ, z et les coordonnées sphériques R, φ, θ dans l'espace euclidien $\mathbf{R}^3$ (fig. 144).

Réponse.

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2\, d\varphi^2 + dz^2 = dR^2 + R^2 \cos^2\theta\, d\varphi^2 + R^2\, d\theta^2$$

Désignons par e_1, e_2, e_3 les vecteurs unitaires du système de coordonnées choisi. Ces vecteurs forment une base dans l'espace tangent en l'origine de ce système.

E x e r c i c e 13. Calculer les formes dx_1, dx_2, dx_3 sur les vecteurs e_1, e_2, e_3.
Réponse. $dx_i(e_i) = 1/\sqrt{E_i}$, les autres formes sont nulles. En particulier,

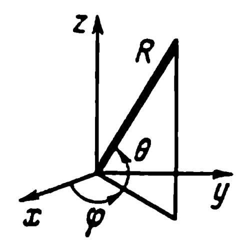


Fig. 144. Problème 12.

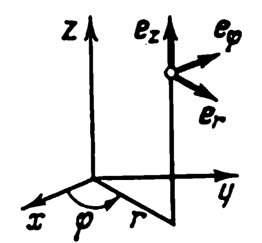


Fig. 145. Problème 13.

en coordonnées cartésiennes $dx(e_x) = dy(e_y) = dz(e_z) = 1$; en coordonnées cylindriques $dr(e_r) = dz(e_z) = 1$, $d\varphi(e_\varphi) = 1/r$ (fig. 145) ; en coordonnées sphériques $dR(e_R) = 1$, $d\varphi(e_\varphi) = \dfrac{1}{R\cos\theta}$, $d\theta(e_\theta) = 1/R$.

La métrique et l'orientation de la variété M confère à l'espace tangent en chaque point de M une structure d'espace euclidien orienté de dimension trois. Au sens de cette structure nous parlerons donc de produits scalaire, vectoriel et mixte.

E x e r c i c e 14. Calculer $[e_1, e_2]$, (e_R, e_θ) et (e_z, e_x, e_y).
Réponse. e_3, 0, 1.

Dans l'espace euclidien orienté $\mathbf{R}^3$, à tout vecteur $\boldsymbol{A}$ correspondent une 1-forme ω_A^1 et une 2-forme ω_A^2 définies par les conditions

$$\omega_A^1(\xi) = (\boldsymbol{A}, \xi), \quad \omega_A^2(\xi, \eta) = (\boldsymbol{A}, \xi, \eta), \quad \forall\, \xi, \quad \eta \in \mathbf{R}^3.$$

La correspondance entre champs de vecteurs et formes ne dépend pas du système de coordonnées mais uniquement de la structure euclidienne et de l'orientation. Donc à chaque champ de vecteurs $\boldsymbol{A}$ sur la variété donnée M correspondent une 1-forme différentielle ω_A^1 sur M et une 2-forme différentielle ω_A^2 sur M.

Les formules de passage des champs aux formes et inversement s'écrivent de façon différente selon le système de coordonnées. Supposons que dans les coordonnées x_1, x_2, x_3 décrites plus haut, le champ de vecteurs $\boldsymbol{A}$ s'écrit

$$\boldsymbol{A} = A_1 e_1 + A_2 e_2 + A_3 e_3$$

(les composantes A_i sont des fonctions différentiables sur M). La 1-forme correspondante ω_A^1 se décompose suivant la base dx_i et la 2-forme ω_A^2 suivant la base $dx_i \wedge dx_j$.

E x e r c i c e 15. Sachant les composantes d'un champ de vecteurs $\boldsymbol{A}$ trouver le développement de la 1-forme ω_A^1 et de la 2-forme ω_A^2.

Solution. On a $\omega_A^1(e_1) = (A, e_1) = A_1$. Par ailleurs, $(a_1\,dx_1 + a_2\,dx_2 + a_3\,dx_3)(e_1) = a_1\,dx_1(e_1) = a_1/\sqrt{E_1}$. D'où il vient $a_1 = A_1\sqrt{E_1}$, de sorte que

$$\omega_A^1 = A_1\sqrt{E_1}\,dx_1 + A_2\sqrt{E_2}\,dx_2 + A_3\sqrt{E_3}\,dx_3.$$

D'une façon analogue on a $\omega_A^2(e_2, e_3) = (A, e_2, e_3) = A_1$. D'autre part

$$(\alpha_1\,dx_2 \wedge dx_3 + \alpha_2\,dx_3 \wedge dx_1 + \alpha_3\,dx_1 \wedge dx_2)(e_2, e_3) = \alpha_1\frac{1}{\sqrt{E_2E_3}}.$$

D'où $\alpha_1 = A_1\sqrt{E_2E_3}$, i.e.

$$\omega_A^2 = A_1\sqrt{E_2E_3}\,dx_2 \wedge dx_3 + A_2\sqrt{E_3E_1}\,dx_3 \wedge dx_1 + A_3\sqrt{E_1E_2}\,dx_1 \wedge dx_2.$$

En particulier, en coordonnées cartésiennes, cylindriques et sphériques sur $\mathbf{R}^3$ au champ de vecteurs

$$A = A_xe_x + A_ye_y + A_ze_z = A_re_r + A_\varphi e_\varphi + A_ze_z = \\ = A_Re_R + A_\varphi e_\varphi + A_\theta e_\theta$$

sont associées la 1-forme

$$\omega_A^1 = A_x\,dx + A_y\,dy + A_z\,dz = A_r\,dr + rA_\varphi d\varphi + A_z\,dz = \\ = A_R\,dR + R\cos\theta A_\varphi d\varphi + RA_\theta\,d\theta$$

et la 2-forme

$$\omega_A^2 = A_x\,dy \wedge dz + A_y\,dz \wedge dx + A_z\,dx \wedge dy = \\ = rA_r\,d\varphi \wedge dz + A_\varphi\,dz \wedge dr + rA_z\,dr \wedge d\varphi = \\ = R^2\cos\theta A_R\,d\varphi \wedge d\theta + RA_\varphi\,d\theta \wedge dR + R\cos\theta A_z\,dR \wedge d\varphi.$$

Un exemple de champ de vecteurs sur une variété M est le gradient d'une fonction $f: M \to \mathbf{R}$. On rappelle que le gradient d'une fonction est le champ de vecteurs **grad** f correspondant à la différentielle :

$$\omega^1_{\mathbf{grad}\,f} = df, \quad \text{i.e.} \quad df(\xi) = (\mathbf{grad}\,f, \xi), \quad \forall\xi.$$

Exercice 16. Trouver les composantes du gradient d'une fonction dans la base e_1, e_2, e_3.

Solution. On a $df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3$.

En vertu de l'exercice précédent

$$\mathbf{grad}\,f = \frac{1}{\sqrt{E_1}}\frac{\partial f}{\partial x_1}e_1 + \frac{1}{\sqrt{E_2}}\frac{\partial f}{\partial x_2}e_2 + \frac{1}{\sqrt{E_3}}\frac{\partial f}{\partial x_3}e_3.$$

En particulier, en coordonnées cartésiennes, cylindriques et sphériques

$$\mathbf{grad}\,f = \frac{\partial f}{\partial x}e_x + \frac{\partial f}{\partial y}e_y + \frac{\partial f}{\partial z}e_z = \frac{\partial f}{\partial r}e_r + \frac{1}{r}\frac{\partial f}{\partial \varphi}e_\varphi + \\ + \frac{\partial f}{\partial z}e_z = \frac{\partial f}{\partial R}e_R = \frac{1}{R\cos\theta}\frac{\partial f}{\partial \varphi}e_\varphi + \frac{1}{R}\frac{\partial f}{\partial \theta}e_\theta.$$

§ 35. Intégration de formes différentielles

On définit ici les notions de chaînes, de bord d'une chaîne et d'intégrale d'une forme étendue à une chaîne.

L'intégrale d'une forme différentielle est la généralisation au cas multidimensionnel de notions telles que le flux d'un fluide à travers une surface ou le travail d'une force pour un déplacement.

A. Intégrale d'une 1-forme étendue à un chemin. Commençons par l'intégration d'une 1-forme ω^1 sur une variété M. Soit $\gamma: [0 \leqslant t \leqslant 1] \to M$ une application différentiable (le « chemin d'intégration »). L'intégrale de la forme ω^1 étendue au chemin γ

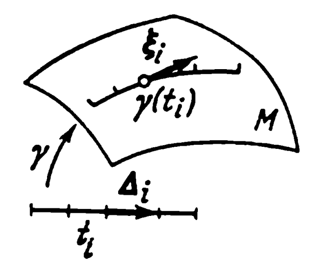


Fig. 146. Intégration d'une 1-forme sur un chemin.

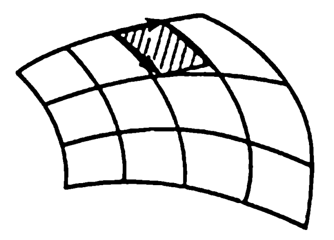


Fig. 147. Intégration d'une 2-forme sur une surface.

se définit comme une limite de sommes intégrales. Chaque somme intégrale est composée de valeurs de la forme ω^1 sur des vecteurs tangents ξ_i (fig. 146):

$$\int_\gamma \omega^1 = \lim_{\Delta \to 0} \sum_{i=1}^{n} \omega^1(\xi_i).$$

Les vecteurs tangents ξ_i se construisent comme suit. On partage l'intervalle $0 \leqslant t \leqslant 1$ en intervalles partiels $\Delta_i: t_i \leqslant t \leqslant t_{i+1}$ par les points t_i. Tout intervalle partiel Δ_i peut être considéré comme un vecteur Δ_i tangent à l'axe t au point t_i. Son image sur l'espace tangent à M au point $\gamma(t_i)$ est

$$\xi_i = d\gamma|_{t_i}(\Delta_i) \in TM_{\gamma(t_i)}.$$

Lorsque le plus grand des intervalles partiels Δ_i tend vers zéro, les sommes intégrales admettent une limite que l'on appelle intégrale de la 1-forme ω^1 étendue au chemin γ.

L'intégrale d'une k-forme étendue à une surface de dimension k se définit de façon analogue. La surface d'intégration est partagée en petits parallélépipèdes curvilignes de dimension k (fig. 147) qui sont remplacés par des parallélépipèdes des espaces tangents, la somme des valeurs de la forme sur les parallélépipèdes des espaces tangents tendant vers l'intégrale lorsque le nombre de subdivisions croît. Voyons d'abord un cas particulier.

B. Intégrale d'une k-forme dans un espace euclidien orienté $\mathbf{R}^k$. Soit $x_1, \ldots, x_k$ un système de coordonnées orienté sur $\mathbf{R}^k$. Toute k-forme sur $\mathbf{R}^k$ est proportionnelle à la forme $dx_1 \wedge \ldots \wedge dx_k$, i.e. est de la forme $\omega^k = \varphi(x)\, dx_1 \wedge \ldots \wedge dx_k$, où $\varphi(x)$ est une fonction différentiable.

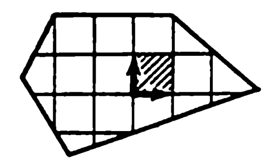


Fig. 148. Intégration d'une k-forme dans un espace de dimension k.

Soit D un polyèdre borné convexe dans $\mathbf{R}^k$ (fig. 148). L'intégrale de la forme ω^k sur D est par définition l'intégrale de la fonction φ:

$$\int_D \omega^k = \int_D \varphi(x)\, dx_1 \ldots dx_k,$$

où l'intégrale du second membre est comprise au sens d'une limite de sommes intégrales riemanniennes ordinaires.

Cette définition va dans le sens du plan tracé plus haut, puisque dans le cas considéré l'espace tangent à la variété s'identifie à cette variété.

Exercice 1. Montrer que $\int_D \omega^k$ dépend linéairement de ω^k.

Exercice 2. Montrer que si on partage D en deux polyèdres convexes D_1 et D_2, alors

$$\int_D \omega^k = \int_{D_1} \omega^k + \int_{D_2} \omega^k.$$

Dans le cas général (une k-forme dans un espace de dimension n) il n'est pas toujours simple d'identifier les éléments de la partition à des parallélépipèdes tangents. Plus bas nous ramènerons ce cas à celui étudié.

C. Effet des applications sur les formes différentielles. Soit $f: M \to N$ une application différentiable d'une variété différentiable

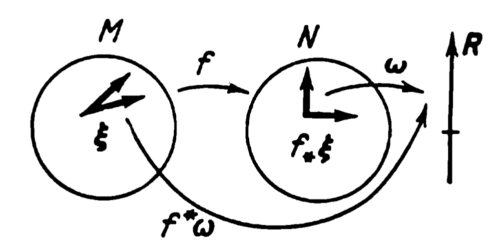


Fig. 149. Une forme sur N induit une forme sur M.

M dans une variété différentiable N et ω une k-forme différentiable sur N (fig. 149).

Alors sur M est également définie une certaine k-forme notée $f^*\omega$ et ayant pour expression

$$(f^*\omega)(\xi_1, \ldots, \xi_k) = \omega(f_*\xi_1, \ldots, f_*\xi_k)$$

quels que soient les vecteurs tangents $\xi_1, \ldots, \xi_k \in TM_x$. Ici f_* représente la différentielle de l'application f.

En d'autres termes la forme $f^*\omega$ prend sur les vecteurs $\xi_1, \ldots$ $\ldots, \xi_k$ une valeur égale à celle de la forme ω sur les images de ces vecteurs.

Exemple. Si $y = f(x_1, x_2) = x_1^2 + x_2^2$ et $\omega = dy$, alors

$$f^*\omega = 2x_1\,dx_1 + 2x_2\,dx_2.$$

Exercice 3. Montrer que $f^*\omega$ est une k-forme sur M.

Exercice 4. Montrer que l'application f^* conserve les opérations sur les formes

$$f^*(\lambda_1\omega_1 + \lambda_2\omega_2) = \lambda_1 f^*(\omega_1) + \lambda_2 f^*(\omega_2),$$
$$f^*(\omega_1 \wedge \omega_2) = (f^*\omega_1) \wedge (f^*\omega_2).$$

Exercice 5. Soit $g: L \to M$ une application différentiable. Montrer que $(fg)^* = g^*f^*$.

Exercice 6. Soient D_1 et D_2 deux polyèdres convexes compacts dans un espace orienté $\mathbf{R}^k$, et $f: D_1 \to D_2$ une application différentiable, dont la restriction à l'intérieur de D_1 est un difféomorphisme de l'intérieur de D_1 sur l'intérieur de D_2 respectant l'orientation. Alors pour toute k-forme différentielle ω^k sur D_2

$$\int_{D_1} f^*\omega^k = \int_{D_2} \omega^k.$$

Indication. C'est le théorème de changement de la variable dans une intégrale multiple:

$$\int_{D_1} \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}\, \varphi(y(x))\, dx_1 \ldots dx_n = \int_{D_2} \varphi(y)\, dy_1 \ldots dy_n.$$

D. Intégration d'une k-forme sur une variété de dimension n. Soit ω une k-forme sur une variété M de dimension n.

Soit D un polyèdre borné convexe de dimension k dans l'espace euclidien $\mathbf{R}^k$ (fig. 150).

Le « chemin d'intégration » sera un morceau σ*) de dimension k défini par le triplet $\sigma = (D, f, \mathrm{Or})$ composé

1) du polyèdre convexe $D \subset \mathbf{R}^k$,

2) de l'application différentiable

$$f: D \to M,$$

3) de l'orientation de $\mathbf{R}^k$ que l'on note Or.

*) D'une façon générale on appelle morceau σ un *polyèdre singulier de dimension k*.

D é f i n i t i o n. On appelle intégrale de la k-forme ω sur le morceau σ de dimension k l'intégrale de la forme correspondante sur le polyèdre D

$$\int_\sigma \omega = \int_D f^*\omega.$$

E x e r c i c e 7. Montrer que l'intégrale dépend linéairement de la forme

$$\int_\sigma \lambda_1\omega_1 + \lambda_2\omega_2 = \lambda_1 \int_\sigma \omega_1 + \lambda_2 \int_\sigma \omega_2.$$

Un morceau de dimension k qui diffère de σ seulement par le choix de l'orientation Or est dit contraire (ou inverse) à σ et se note $-\sigma$ ou $-1\cdot\sigma$ (fig. 151).

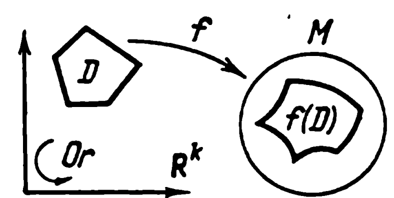



Fig. 150. k-polyèdre singulier.

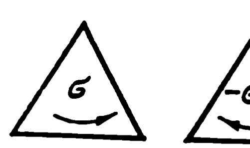


Fig. 151. Problème 8.

E x e r c i c e 8. Montrer que lorsqu'on change l'orientation, l'intégrale change de signe:

$$\int_{-\sigma} \omega = -\int_\sigma \omega.$$

E. Chaînes. L'ensemble $f(D)$ n'est pas forcément une sous-variété différentiable de M. Il peut présenter des points « doubles », des « plis» de toute nature et même dégénérer en un point. Cependant il est clair que même dans le cas unidimensionnel il n'est pas commode de se borner à des contours d'intégration composés d'un morceau seulement: il y a également intérêt à considérer des contours constitués de plusieurs morceaux susceptibles d'être parcourus plusieurs fois dans les deux sens. Dans le cas multidimensionnel cette notion s'appelle *chaîne*.

D é f i n i t i o n. Une *chaîne de dimension k* (ou simplement une k-chaîne) *sur une variété M* est composée d'une collection finie de morceaux orientés $\sigma_1, \ldots, \sigma_r$ de dimension k sur M et de nombres entiers $m_1, \ldots, m_r$ appelés *multiplicités* (les multiplicités peuvent être positives, négatives ou nulles). La chaîne est désignée par

$$c_k = m_1\sigma_1 + \ldots + m_r\sigma_r.$$

Ceci étant, on introduit les identités

$$m_1\sigma + m_2\sigma = (m_1 + m_2)\,\sigma,$$
$$m_1\sigma_1 + m_2\sigma_2 = m_2\sigma_2 + m_1\sigma_1, \quad 0\sigma = 0, \quad c_k + 0 = c_k.$$

E x e r c i c e 9. Démontrer que les chaînes de dimension k sur M forment un groupe commutatif si l'on définit la somme de chaînes par la formule

$$(m_1\sigma_1 + \ldots + m_r\sigma_r) + (m'_1\sigma'_1 + \ldots + m'_{r'}\sigma'_{r'}) =$$
$$= m_1\sigma_1 + \ldots + m_r\sigma_r + m'_1\sigma'_1 + \ldots + m'_{r'}\sigma'_{r'}.$$

F. Exemple: bord d'un polyèdre. Soit D un polyèdre convexe orienté de dimension k dans l'espace euclidien $\mathbf{R}^k$. Le *bord* de D est par définition une $(k-1)$-chaîne ∂D sur $\mathbf{R}^k$ définie comme suit (fig. 152).

Les morceaux σ_i de la chaîne ∂D sont composés des faces D_i de dimension $k-1$ du polyèdre D, des applications $f_i\colon D_i \to \mathbf{R}^k$

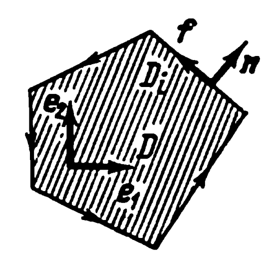


Fig. 152. Orientation du bord.

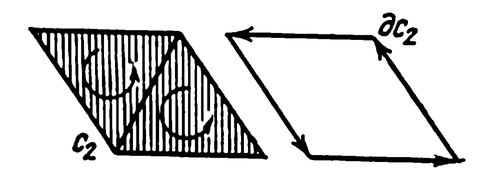


Fig. 153. Bord d'une chaîne.

qui plongent les faces D_i dans $\mathbf{R}^k$ et des orientations Or_i définies plus bas; les multiplicités sont égales à 1 :

$$\partial D = \sum \sigma_i, \quad \sigma_i = (D_i, f_i, \mathrm{Or}_i).$$

R è g l e d' o r i e n t a t i o n d e s f a c e s. Soient $e_1, \ldots, e_k$ un repère orienté dans $\mathbf{R}^k$, D_i une face de D. En un point situé à l'intérieur de D_i traçons le vecteur n de la normale extérieure au polyèdre D. Un *repère orientant la face* D_i est un repère $f_1, \ldots, f_{k-1}$ tel que le repère $(n, f_1, \ldots, f_{k-1})$ soit orienté *correctement* (i.e. comme $e_1, \ldots, e_k$).

Le *bord d'une chaîne* se définit de façon analogue. Soit $\sigma = (D, f, \mathrm{Or})$ un morceau de dimension k sur M. Son bord $\partial\sigma$ est une $(k-1)$-chaîne $\partial\sigma = \sum \sigma_i$ composée des morceaux $\sigma_i = (D_i, f_i, \mathrm{Or}_i)$, où D_i sont les $(k-1)$-faces de D, Or_i les orientations choisies d'après la règle indiquée, f_i la restriction de l'application $f\colon D \to M$ à la face D_i.

On appelle bord ∂c_k d'une chaîne c_k de dimension k sur M la somme des bords des morceaux de la chaîne c_k avec leurs multiplicités (fig. 153):

$$\partial c_k = \partial (m_1\sigma_1 + \ldots + m_r\sigma_r) = m_1\partial\sigma_1 + \ldots + m_r\partial\sigma_r.$$

De toute évidence ∂c_k est une $(k-1)$-chaîne sur M *).

E x e r c i c e 10. Montrer que le bord d'une chaîne quelconque est nul: $\partial\partial c_k = 0$.

I n d i c a t i o n. ∂ étant linéaire, il suffit de montrer que $\partial\partial D = 0$ pour le polyèdre convexe D. Il reste à vérifier que toute $(k-2)$-face de D figure deux fois dans la chaîne $\partial\partial D$ avec des signes différents. Il suffit de le vérifier pour $k = 2$ (les sections planes).

F. Intégrale d'une forme étendue à une chaîne. Soit ω^k une k-forme sur une variété M et c_k une k-chaîne sur M, $c_k = \sum m_i\sigma_i$. On appelle *intégrale de la forme* ω^k *étendue à la chaîne* c_k la somme des intégrales étendues aux morceaux σ_i compte tenu des multiplicités

$$\int_{c_k} \omega^k = \sum m_i \int_{\sigma_i} \omega^k.$$

E x e r c i c e 11. Montrer que l'intégrale dépend linéairement de la forme

$$\int_{c_k} \omega_1^k + \omega_2^k = \int_{c_k} \omega_1^k + \int_{c_k} \omega_2^k.$$

E x e r c i c e 12. Montrer que l'intégration d'une forme ω^k fixe étendue aux chaînes c_k est un homomorphisme du groupe des chaînes sur une droite.

E x e m p l e 1. Supposons que M est un plan $\{(p, q)\}$, $p\,dq$ une forme ω^1 et c_1 une chaîne composée d'un seul morceau σ_1 de multiplicité 1:

$$[0 \leqslant t \leqslant 2\pi] \xrightarrow{f} (p = \cos t,\ q = \sin t).$$

Donc

$$\int_{c_1} p\,dq = \pi.$$

D'une façon générale, si la chaîne c_1 est le bord d'un domaine G (fig. 154), alors $\int_{c_1} p\,dq$ est égale à l'aire du domaine G prise avec le signe + ou — selon que le couple de vecteurs composé de la normale extérieure et du vecteur orientant le bord a la même orientation ou non que le repère auquel est rapporté le plan.

E x e m p l e 2. Supposons que M est l'espace euclidien orienté $\mathbf{R}^3$. Alors toute 1-forme sur M est associée à un champ de vecteurs A $(\omega^1 = \omega_A^1)$, où

$$\omega_A^1(\xi) = (A, \xi).$$

On appelle *circulation du champ A le long d'une courbe orientée l* l'intégrale de la forme ω_A^1 étendue à une chaîne c_1 assimilée à la courbe l:

$$\int_{c_1} \omega_A^1 = \int_l (A, dl).$$

*) On suppose que $k > 1$. Les chaînes de dimension 1 font partie du schéma général si l'on introduit la définition suivante : une chaîne de dimension nulle est constituée d'une collection de points avec leurs multiplicités; le bord d'un segment de droite orienté $\overrightarrow{AB}$ est $B - A$ (le point B est de multiplicité 1, le point A de multiplicité -1); le bord d'un point est vide.

Toute 2-forme sur M est associée également à un champ $\boldsymbol{A}$ ($\omega^2 = \omega^2_{\boldsymbol{A}}$, où $\omega^2_{\boldsymbol{A}}(\xi, \eta) = (\boldsymbol{A}, \xi, \eta)$).

On appelle *flux du champ* $\boldsymbol{A}$ *à travers une surface orientée* S l'intégrale de la forme $\omega^2_{\boldsymbol{A}}$ étendue à une chaîne c_2 assimilée à S:

$$\int_{c_2} \omega^2_{\boldsymbol{A}} = \int_S (\boldsymbol{A}, d\boldsymbol{n}).$$

Exercice 13. Trouver le flux du champ $\boldsymbol{A} = \frac{1}{R^2} e_R$ à travers la surface de la sphère $x^2 + y^2 + z^2 = 1$ orientée par le repère e_x, e_y au point $z = 1$.

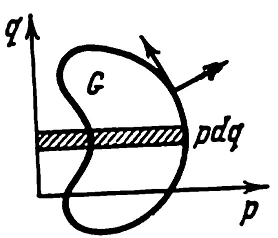


Fig. 154. L'intégrale de la forme $p\,dq$ sur le bord du domaine est égale à l'aire de ce domaine.

Trouver le flux du même champ à travers la surface de l'ellipsoïde $x^2/a^2 + y^2/b^2 + z^2 = 1$ orientée de la même façon.

Indication. Voir la fin du chapitre 7.

Exercice 14. Dans un espace $\mathbf{R}^{2n} = \{(p_1, \ldots, p_n; q_1, \ldots, q_n)\}$ on donne une 2-chaîne c_2 assimilée à une surface orientée S de dimension 2 et de bord l. Calculer

$$\int_{c_2} dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n \quad \text{et} \quad \int_l p_1\,dq_1 + \ldots + p_n\,dq_n.$$

Réponse. La somme des aires orientées des projections de S sur les plans arithmétiques p_i, q_i de dimension 2.

§ 36. Différentiation extérieure

On définit ici la différentiation extérieure d'une k-forme et on démontre la formule de Stokes: l'intégrale de la dérivée d'une forme étendue à une chaîne est égale à l'intégrale de cette forme prise le long du bord de cette chaîne.

A. Exemple: divergence d'un champ de vecteurs. La dérivée extérieure d'une k-forme ω sur une variété M est une $(k+1)$-forme $\partial\omega$ sur la même variété. On passe d'une forme à sa dérivée extérieure de la même manière qu'on prend la différentielle d'une fonction ou la divergence d'un champ de vecteurs. Rappelons la définition de la divergence.

Soient A un champ de vecteurs sur l'espace euclidien orienté $\mathbf{R}^3$ et S le bord d'un parallélépipède P d'arêtes ξ_1, ξ_2, ξ_3 et de sommet x: $S = \partial P$ (fig. 155). Etudions le flux du champ $\boldsymbol{A}$ à travers S (« vers l'extérieur »):

$$F(P) = \int_S (\boldsymbol{A}, d\boldsymbol{n}).$$

Si le parallélépipède est très petit, le flux F est approximativement proportionnel au produit du volume du parallélépipède

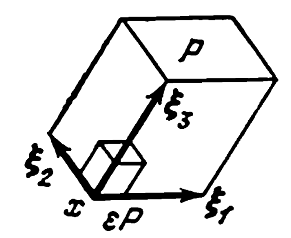


Fig. 155. Divergence d'un champ de vecteurs.

$V = (\xi_1, \xi_2, \xi_3)$ par la « densité des sources » au point x. En d'autres termes, existe la limite

$$\lim_{\varepsilon \to 0} \frac{F(\varepsilon P)}{\varepsilon^3 V},$$

où εP est un parallélépipède d'arêtes $\varepsilon\xi_1$, $\varepsilon\xi_2$, $\varepsilon\xi_3$. Cette limite ne dépend pas du choix du parallélépipède P, mais seulement du point x. On l'appelle *divergence du champ* A en x et on la note div A.

Pour passer au cas multidimensionnel remarquons que le « flux A à travers l'élément de surface » est une 2-forme que nous avons désignée par ω_A^2. La divergence n'est autre que la densité dans l'expression de la 3-forme

$$\omega^3 = \operatorname{div} A \, dx \wedge dy \wedge dz,$$

$$\omega^3 (\xi_1, \xi_2, \xi_3) = \operatorname{div} A \cdot V (\xi_1, \xi_2, \xi_3),$$

caractérisant les « sources dans un parallélépipède élémentaire ».

La dérivée extérieure $d\omega^k$ de la k-forme ω^k sur M se définit comme la partie multilinéaire principale de l'intégrale de ω^k étendue au bord d'un parallélépipède de dimension $k+1$.

B. Définition de la dérivée extérieure. Définissons la valeur de la forme $d\omega$ sur $k+1$ vecteurs $\xi_1, \ldots, \xi_{k+1}$ tangents à M en x.

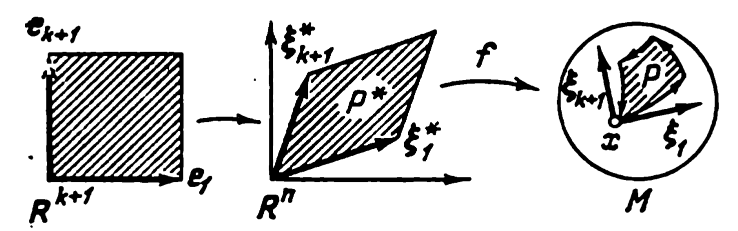


Fig. 156. Parallélépipède curviligne P.

Considérons pour cela un système de coordonnées quelconque au voisinage du point x sur M, i. e. un difféomorphisme f d'un voisinage du point 0 de l'espace euclidien $\mathbf{R}^n$ sur un voisinage du point x sur M (fig. 156).

Les antécédents des vecteurs $\xi_1, \ldots, \xi_{k+1} \in TM_x$ par la différentielle de f sont situés dans l'espace tangent à $\mathbf{R}^n$ en 0. Cet espace tangent s'identifie naturellement à $\mathbf{R}^n$, c'est pourquoi on peut considérer que les antécédents sont les vecteurs

$$\xi_1^*, \ldots, \xi_{k+1}^* \in \mathbf{R}^n.$$

Construisons sur ces vecteurs de $\mathbf{R}^n$ un parallélépipède P^* (en toute rigueur il faudrait considérer un cube standard orienté dans $\mathbf{R}^{k+1}$ et son application linéaire sur P^* qui aux arêtes $e_1, \ldots, e_{k+1}$ associe $\xi_1^*, \ldots, \xi_{k+1}^*$ en tant que $(k+1)$-morceau dans $\mathbf{R}^n$).

L'application f associe au parallélépipède P^* un $(k+1)$-morceau P sur M (« un parallélépipède curviligne »). Le bord du morceau P est une k-chaîne ∂P. Considérons l'intégrale d'une forme ω^k le long du bord ∂P du parallélépipède P:

$$F(\xi_1, \ldots, \xi_{k+1}) = \int_{\partial P} \omega^k.$$

Exemple. Nous appellerons 0-forme sur M une fonction différentiable $\varphi: M \to \mathbf{R}$. Nous appellerons intégrale de la 0-forme φ étendue à une 0-chaîne $c_0 = \Sigma m_i A_i$ (où m_i sont des entiers et A_i des points de M) l'expression

$$\int_{c_0} \varphi = \sum m_i \varphi(A_i).$$

Alors la définition précédente nous donne l'« accroissement » $F(\xi_1) = \varphi(x_1) - \varphi(x)$ de la fonction (fig. 157); quant à la partie linéaire principale de $F(\xi_1)$ en 0, c'est tout simplement la différentielle de la fonction φ.

Exercice 1. Montrer que la fonction $F(\xi_1, \ldots, \xi_{k+1})$ est antisymétrique en ξ.

Il s'avère que la $(k+1)$-partie linéaire principale de l'« accroissement » de $F(\xi_1, \ldots, \xi_{k+1})$ est une $(k+1)$-forme extérieure

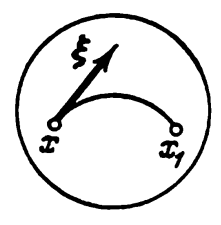


Fig. 157. L'intégrale sur le bord d'un 1-parallélépipède est égale à l'accroissement d'une fonction.

sur l'espace TM_x tangent à M en x. Cette forme ne dépend pas du système de coordonnées qui a servi à définir le parallélépipède curviligne P. On l'appelle *dérivée extérieure* de la forme ω^k (au point x) et on la note $d\omega^k$.

C. Théorème de la dérivée extérieure.

Théorème. *Sur TM_x il existe une $(k+1)$-forme Ω et une seule, $(k+1)$-partie linéaire principale en 0 de l'intégrale $F(\xi_1, \ldots, \xi_{k+1})$,*

étendue au bord d'un parallélépipède curviligne, i.e.

$$F(\varepsilon\xi_1, \ldots, \varepsilon\xi_{k+1}) = \varepsilon^{k+1}\Omega(\xi_1, \ldots, \xi_{k+1}) + o(\varepsilon^{k+1}) \quad (\varepsilon \to 0). \quad (1)$$

La forme Ω ne dépend pas du choix du système de coordonnées qui a participé à la définition de F.

Si dans un système local de coordonnées $x_1, \ldots, x_n$ sur M la forme ω^k s'écrit

$$\omega^k = \sum a_{i_1 \ldots i_k}\, dx_{i_1} \wedge \ldots \wedge dx_{i_k},$$

la forme Ω aura pour expression

$$\Omega = d\omega^k = \sum da_{i_1 \ldots i_k} \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k}. \quad (2)$$

Nous démontrerons ce théorème pour le cas de la forme $\omega^1 = a(x_1, x_2)\, dx_1$ sur le plan (x_1, x_2). Dans le cas général la démonstration est exactement la même, mais un peu plus longue.

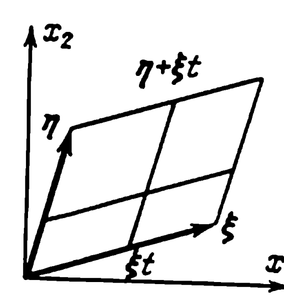


Fig. 158. Sur le théorème relatif à la dérivée extérieure.

Calculons F, i. e. l'intégrale de ω^1 étendue au bord d'un parallélogramme P de côtés ξ, η et de sommet 0 (fig. 158). La chaîne ∂P est définie par les applications $t \to \xi t$, $t \to \xi + \eta t$, $t \to \eta t$, $t \to \eta + \xi t$ respectivement de multiplicités 1, 1, -1, -1 de l'intervalle $0 \leqslant t \leqslant 1$ sur le plan. Donc

$$\int_{\partial P} \omega^1 = \int_0^1 [a(\xi t) - a(\xi t + \eta)]\, \xi_1 - [a(\eta t) - a(\eta t + \xi)]\, \eta_1\, dt,$$

où $\xi_1 = dx_1(\xi)$, $\eta_1 = dx_1(\eta)$, $\xi_2 = dx_2(\xi)$, $\eta_2 = dx_2(\eta)$ sont les composantes des vecteurs ξ, η. Or

$$a(\xi t + \eta) - a(\xi t) = \frac{\partial a}{\partial x_1}\eta_1 + \frac{\partial a}{\partial x_2}\eta_2 + O(\xi^2, \eta^2)$$

(la dérivation s'effectue pour $x_1 = x_2 = 0$). De la même façon

$$a(\eta t + \xi) - a(\eta t) = \frac{\partial a}{\partial x_1}\xi_1 + \frac{\partial a}{\partial x_2}\xi_2 + O(\xi^2, \eta^2).$$

En portant ces expressions dans l'intégrale on trouve

$$F(\xi, \eta) = \int_{\partial P} \omega^1 = \frac{\partial a}{\partial x_2}(\xi_2\eta_1 - \xi_1\eta_2) + o(\xi^2, \eta^2).$$

On remarquera que la partie bilinéaire principale de F, comme nous l'avons annoncé dans (1), est la valeur de la 2-forme

$$\Omega = \frac{\partial a}{\partial x_2} dx_2 \wedge dx_1$$

sur le couple de vecteurs ξ, η. De plus la forme obtenue est donnée par la formule (2), puisque

$$da \wedge dx_1 = \frac{\partial a}{\partial x_1} dx_1 \wedge dx_1 + \frac{\partial a}{\partial x_2} dx_2 \wedge dx_1 = \frac{\partial a}{\partial x_2} dx_2 \wedge dx_1.$$

Enfin si l'on remplace le système de coordonnées x_1, x_2 par un autre (fig. 159), alors le parallélogramme P est remplacé par un parallé-

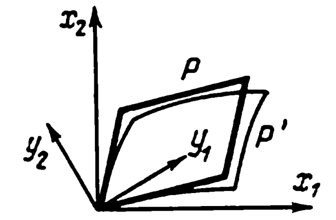


Fig. 159. Indépendance de la dérivée extérieure par rapport au système de coordonnées.

logramme curviligne P' proche, de sorte que la différence des intégrales $\int_{\partial P} \omega^1 - \int_{\partial P'} \omega^1$ sera un infiniment petit d'ordre supérieur au second (vérifier !), c.q.f.d.

Exercice 2. Faire la démonstration dans le cas général.

Exercice 3. Démontrer les formules de différentiation de la somme et du produit:

$$d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2,$$
$$d(\omega^k \wedge \omega^l) = d\omega^k \wedge \omega^l + (-1)^k \omega^k \wedge d\omega^l.$$

Exercice 4. Montrer que la différentielle d'une différentielle est nulle: $dd = 0$.

Exercice 5. Soit $f: M \to N$ une application différentiable, ω une k-forme sur N. Montrer que $f^* (d\omega) = d(f^*\omega)$.

D. Formule de Stokes. L'une des plus importantes conséquences du théorème de la dérivée extérieure est la formule de Newton-Leibniz-Gauss-Green-Ostrogradski-Stokes-Poincaré:

$$\int_{\partial c} \omega = \int_c d\omega, \tag{3}$$

où c est une $(k+1)$-chaîne quelconque sur une variété M, ω une k-forme quelconque sur M.

Il suffit de démontrer cette formule pour le cas où la chaîne est composée d'un seul morceau σ. Supposons d'abord que ce morceau σ est défini par un parallélépipède orienté $P \subset \mathbf{R}^{k+1}$ (fig. 160).

Divisons P en N^{k+1} petits parallélépipèdes égaux P_i semblables à P. Alors de toute évidence

$$\int_{\partial P} \omega = \sum_{i=1}^{N^{k+1}} F_i, \quad \text{où} \quad F_i = \int_{\partial P_i} \omega.$$

D'après la formule (1) on a

$$F_i = d\omega\,(\xi_1^i, \ldots, \xi_{k+1}^i) + o\,(N^{-(k+1)}),$$

où $\xi_1^i, \ldots, \xi_{k+1}^i$ sont les arêtes de P_i. Or $\sum_{i=1}^{N^{k+1}} d\omega\,(\xi_1^i, \ldots, \xi_{k+1}^i)$

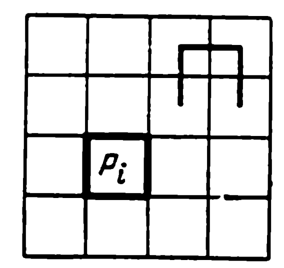


Fig. 160. Démonstration de la formule de Stokes pour le parallélépipède.

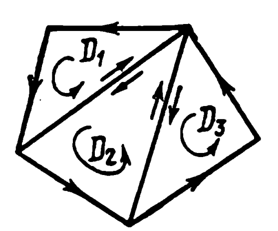


Fig. 161. Partition d'un polyèdre convexe en simplexes.

est somme intégrale de $\int_P d\omega$. Il est aisé de vérifier que $o\,(N^{-(k+1)})$ est uniforme, donc

$$\lim_{N\to\infty} \sum_{i=1}^{N^{k+1}} F_i = \lim_{N\to\infty} \sum_{i=1}^{N^{k+1}} d\omega\,(\xi_1^i, \ldots, \xi_{k+1}^i) = \int_P d\omega.$$

Finalement on obtient

$$\int_{\partial P} \omega = \sum F_i = \lim_{N\to\infty} \sum F_i = \int_P d\omega.$$

Ceci entraîne aussitôt la formule (3) pour toute chaîne dont les polyèdres sont des parallélépipèdes.

Pour démontrer la formule (3) pour un polyèdre convexe quelconque D, il suffit de la démontrer pour un simplexe *), puisque D

*) Un simplexe de dimension 2 est un triangle, celui de dimension 3, un tétraèdre ; de dimension k, enveloppe convexe de $(k+1)$ points dans $\mathbf{R}^n$ non situés dans un plan de dimension $k-1$.

Exemple. $\left\{x \in \mathbf{R}^k\colon x_i \geqslant 0, \ \sum_{i=1}^{k} x_i \leqslant 1\right\}$.

peut toujours être décomposé en simplexes (fig. 161):

$$D = \sum D_i, \quad \partial D = \sum \partial D_i.$$

Démontrons la formule (3) pour un simplexe. Remarquons qu'il existe une application différentiable d'un k-cube orienté sur un k-simplexe orienté dont

1) la restriction à l'intérieur du cube est un difféomorphisme de l'intérieur du cube sur l'intérieur du simplexe respectant l'orientation;

2) la restriction à certaines $(k-1)$-faces du cube est un difféomorphisme de l'intérieur des faces du cube sur l'intérieur des faces du simplexe respectant l'orientation; les images des autres $(k-1)$-faces du cube sont contenues dans des $(k-2)$-faces du simplexe.

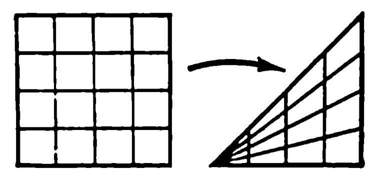


Fig. 162. Démonstration de la formule de Stokes pour le simplexe.

Pour $k = 2$, par exemple, une telle application du cube $0 \leqslant x_1, x_2 \leqslant 1$ sur un triangle est donnée par les formules $y_1 = x_1$, $y_2 = x_1x_2$ (fig. 162). Et la formule (3) pour le simplexe découle de la formule (3) démontrée pour le cube et du théorème de changement de variables (voir § 35, p. C).

E x e m p l e 1. Considérons dans $\mathbf{R}^{2n}$ muni des coordonnées $p_1, \ldots, p_n; q_1, \ldots, q_n$ la 1-forme

$$\omega^1 = p_1\, dq_1 + \ldots + p_n\, dq_n = \boldsymbol{p}\, d\boldsymbol{q}.$$

Alors $d\omega^1 = dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n = d\boldsymbol{p} \wedge d\boldsymbol{q}$, donc

$$\iint_{c_2} d\boldsymbol{p} \wedge d\boldsymbol{q} = \int_{\partial c_2} \boldsymbol{p}\, d\boldsymbol{q}.$$

En particulier, si c_2 est une surface fermée $(\partial c_2 = 0)$, alors $\iint_{c_2} d\boldsymbol{p} \wedge d\boldsymbol{q} = 0$.

E. Exemple 2. Analyse vectorielle. Dans un espace orienté riemannien M de dimension trois, à tout champ de vecteurs $\boldsymbol{A}$ sont associées une 1-forme $\omega^1_{\boldsymbol{A}}$ et une 2-forme $\omega^2_{\boldsymbol{A}}$. Donc la différentiation extérieure peut être envisagée comme une opération vectorielle.

A la différentiation extérieure des 0-formes (des fonctions), des 1-formes et des 2-formes correspondent les opérations gradient, rotationnel et divergence définies par les expressions

$$df = \omega^1_{\operatorname{grad} f}, \quad d\omega^1_{\boldsymbol{A}} = \omega^2_{\operatorname{rot} \boldsymbol{A}}, \quad \partial\omega^2_{\boldsymbol{A}} = \operatorname{div} \boldsymbol{A}\omega^3$$

(la forme ω^3 est l'élément de volume sur M). Donc de (3) il vient

$$f(y)-f(x)=\int_l \operatorname{grad} f\cdot dl \quad \text{si} \quad \partial l = y-x,$$

$$\int_l A\,dl=\iint_S \operatorname{rot} A\cdot dn \quad \text{si} \quad dS=l,$$

$$\iint_S A\,dn=\iiint_D \operatorname{div} A\omega^3 \quad \text{si} \quad \partial D=S.$$

Exercice 5. Montrer que

$$\operatorname{div}[A, B]=(\operatorname{rot} A, B)-(\operatorname{rot} B, A),$$
$$\operatorname{rot} aA=[\operatorname{grad} a, A]+a\operatorname{rot} A,$$
$$\operatorname{div} aA=(\operatorname{grad} a, A)+a\operatorname{div} A.$$

Indication. D'après la formule de différentiation du produit de formes

$$d(\omega^2_{[A,B]})=d(\omega^1_A\wedge\omega^1_B)=d\omega^1_A\wedge\omega^1_B-\omega^1_A\wedge d\omega^1_B.$$

Exercice 6. Montrer que rot grad = div rot = 0.
Indication. $dd=0$.

F. Appendice 1. Opérations vectorielles dans un système trirectangulaire. Soient x_1, x_2, x_3 un système de coordonnées trirectangulaire dans M, $ds^2=E_1\,dx_1^2+E_2\,dx_2^2+E_3\,dx_3^2$ et e_i les vecteurs unitaires (voir § 34, p. F).

Exercice 8. Sachant les composantes du champ de vecteurs $A = A_1e_1+A_2e_2+A_3e_3$, trouver celles de son rotationnel.

Solution. En vertu de l'exercice 15, § 34, on a

$$\omega^1_A=A_1\sqrt{E_1}\,dx_1+A_2\sqrt{E_2}\,dx_2+A_3\sqrt{E_3}\,dx_3.$$

Donc

$$d\omega^1_A=\left(\frac{\partial A_3\sqrt{E_3}}{\partial x_2}-\frac{\partial A_2\sqrt{E_2}}{\partial x_3}\right)dx_2\wedge dx_3+\ldots=\omega^2_{\operatorname{rot} A}.$$

Il vient de l'exercice 15

$$\operatorname{rot} A=\frac{1}{\sqrt{E_2E_3}}\left(\frac{\partial A_3\sqrt{E_3}}{\partial x_2}-\frac{\partial A_2\sqrt{E_2}}{\partial x_3}\right)e_1+\ldots$$

$$\ldots=\frac{1}{\sqrt{E_1E_2E_3}}\begin{vmatrix}\sqrt{E_1}\,e_1 & \sqrt{E_2}\,e_2 & \sqrt{E_3}\,e_3\\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3}\\ A_1\sqrt{E_1} & A_2\sqrt{E_2} & A_3\sqrt{E_3}\end{vmatrix}.$$

En particulier, en coordonnées cartésiennes, cylindriques et sphériques dans $\mathbf{R}^3$:

$$\operatorname{rot} A = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) e_x + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) e_y + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) e_z =$$

$$= \frac{1}{r}\left(\frac{\partial A_z}{\partial \varphi} - \frac{\partial r A_\varphi}{\partial z}\right) e_r + \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}\right) e_\varphi + \frac{1}{r}\left(\frac{\partial r A_\varphi}{\partial r} - \frac{\partial A_r}{\partial \varphi}\right) e_z =$$

$$= \frac{1}{R\cos\theta}\left(\frac{\partial A_\theta}{\partial \varphi} - \frac{\partial A_\varphi \cos\theta}{\partial \theta}\right) e_R + \frac{1}{R}\left(\frac{\partial A_R}{\partial \theta} - \frac{\partial R A_\theta}{\partial R}\right) e_\varphi +$$

$$+ \frac{1}{R}\left(\frac{\partial R A_\varphi}{\partial R} - \frac{1}{\cos\theta}\frac{\partial A_R}{\partial \varphi}\right) e_\theta.$$

Exercice 9. Trouver la divergence du champ $A = A_1 e_1 + A_2 e_2 + A_3 e_3$.
Solution. $\omega_A^2 = A_1 \sqrt{E_2 E_3}\, dx_2 \wedge dx_3 + \ldots$. Donc,

$$d\omega_A^2 = \frac{\partial}{\partial x_1}(A_1 \sqrt{E_2 E_3})\, dx_1 \wedge dx_2 \wedge dx_3 + \ldots .$$

Selon la définition de la divergence

$$d\omega_A^2 = \operatorname{div} A \sqrt{E_1 E_2 E_3}\, dx_1 \wedge dx_2 \wedge dx_3.$$

Il en résulte

$$\operatorname{div} A = \frac{1}{\sqrt{E_1 E_2 E_3}}\left(\frac{\partial}{\partial x_1} A_1 \sqrt{E_2 E_3} + \frac{\partial}{\partial x_2} A_2 \sqrt{E_3 E_1} + \frac{\partial}{\partial x_3} A_3 \sqrt{E_1 E_2}\right).$$

En particulier, en coordonnées cartésiennes, cylindriques et sphériques dans $\mathbf{R}^3$

$$\operatorname{div} A = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \frac{1}{r}\left(\frac{\partial r A_r}{\partial r} + \frac{\partial A_\varphi}{\partial \varphi}\right) + \frac{\partial A_z}{\partial z} =$$

$$= \frac{1}{R^2\cos\theta}\left(\frac{\partial R^2 \cos\theta A_R}{\partial R} + \frac{\partial R A_\varphi}{\partial \varphi} + \frac{\partial R\cos\theta A_z}{\partial z}\right).$$

Exercice 10. On appelle opérateur de Laplace sur M l'opérateur $\Delta = \operatorname{div}\operatorname{grad}$. Trouver son expression dans les coordonnées x_i.
Réponse.

$$\Delta f = \frac{1}{\sqrt{E_1 E_2 E_3}}\left[\frac{\partial}{\partial x_1}\left(\sqrt{\frac{E_2 E_3}{E_1}}\frac{\partial f}{\partial x_1}\right) + \ldots\right].$$

En particulier, dans $\mathbf{R}^3$:

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r}\frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2} =$$

$$= \frac{1}{R^2\cos\theta}\left[\frac{\partial}{\partial R}\left(R^2\cos\theta\frac{\partial f}{\partial R}\right) + \frac{\partial}{\partial \varphi}\left(\frac{1}{\cos\theta}\frac{\partial}{\partial \varphi}\right) + \frac{\partial}{\partial \theta}\left(\cos\theta\frac{\partial f}{\partial \theta}\right)\right].$$

F. Appendice 2. Formes fermées et cycles. Le flux d'un fluide incompressible (sans sources) à travers le bord d'un domaine D est nul. Généralisons cette proposition évidente au cas multidimensionnel.

L'équivalent multidimensionnel d'un flux sans sources s'appelle *forme fermée*.

Un champ A ne possède pas de sources si div $A = 0$.

D é f i n i t i o n. Une forme différentielle ω sur une variété M est *fermée* si sa dérivée extérieure est nulle : $d\omega = 0$.

En particulier, une 2-forme ω_A^2 associée à un champ de vecteurs sans sources A est fermée. La formule de Stokes (3) entraîne immédiatement le

Théorème. *L'intégrale d'une forme fermée ω^k le long du bord d'une $(k+1)$-chaîne quelconque c_{k+1} est nulle*:

$$\int_{\partial c_{k+1}} \omega^k = 0 \quad si \quad d\omega^k = 0.$$

E x e r c i c e 11. Montrer que la différentielle d'une forme est toujours fermée.

La réciproque n'est pas vraie : il existe en effet des formes fermées qui ne sont pas différentielles.

Prenons par exemple pour variété M l'espace euclidien $\mathbf{R}^3$ privé du point O: $M = \mathbf{R}^3 - \{O\}$ et, pour 2-forme, le flux du champ $A = \frac{1}{R^2} e_R$ (fig. 163). On s'assure immédiatement que div $A = 0$,

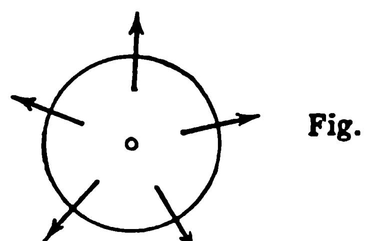


Fig. 163. Champ A.

de sorte que notre 2-forme ω_A^2 est fermée. D'autre part, le flux à travers une sphère quelconque de centre O est égal à 4π. Montrons que l'intégrale de la différentielle de la forme sur la sphère est nulle.

D é f i n i t i o n. On appelle *cycle* sur une variété M une chaîne dont le bord est nul.

La surface orientée de notre sphère peut être considérée comme un cycle. La formule de Stokes (3) entraîne aussitôt le

Théorème. *L'intégrale d'une différentielle sur un cycle quelconque est nulle*:

$$\int_{c_{k+1}} d\omega^k = 0 \quad si \quad \partial c_{k+1} = 0.$$

Donc, notre 2-forme ω_A^2 n'est différentielle d'aucune 1-forme.

L'existence sur M de formes fermées qui ne sont pas différentielles tient aux propriétés topologiques de M. On montre que toute

k-forme fermée dans un espace vectoriel est différentielle d'une $(k+1)$-forme (« lemme de Poincaré »).

E x e r c i c e 12. Démontrer le lemme de Poincaré pour les 1-formes.

I n d i c a t i o n. Etudier $\int_{x_0}^{x_1} \omega^1 = \varphi(x_1)$.

E x e r c i c e 13. Démontrer que sur un espace vectoriel l'intégrale d'une forme fermée sur un cycle quelconque est nulle.

I n d i c a t i o n. Construire une $(k+1)$-chaîne dont le bord est le cycle donné (fig. 164).

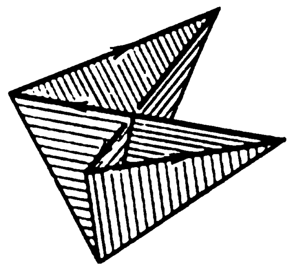


Fig. 164. Cône au-dessus d'un cycle.

Plus exactement, considérer pour toute chaîne c le « cône de sommet 0 construit sur c ». Si l'on désigne par p l'opération de construction du cône, on a

$$\partial \circ p + p \circ \partial = 1 \text{ (transformation identique).}$$

Donc si la chaîne c est fermée, $\partial(pc) = c$.

E x e r c i c e. Démontrer que toute forme fermée dans un espace vectoriel est une différentielle totale.

I n d i c a t i o n. Se servir de la construction précédente. Soit ω^k une k-forme différentielle sur $\mathbf{R}^n$. Définissons une $(k-1)$-forme différentielle $p\omega^k$ (« cocône sur ω ») comme suit: pour toute chaîne c_{k-1}

$$\int_{c_{k-1}} p\omega^k = \int_{pc_k} \omega^k.$$

On vérifie aisément que la $(k-1)$-forme $p\omega^k$ existe et est unique; sur des vecteurs $\xi_1, \ldots, \xi_{k-1}$ tangents à $\mathbf{R}^n$ en x, elle prend la valeur

$$(p\omega)_x(\xi_1, \ldots, \xi_{k-1}) = \int_0^1 \omega_{t_x}(x, t\xi_1 \ldots, t\xi_{k-1})\, dt.$$

Il est aisé de vérifier que

$$d \circ p + p \circ d = 1 \text{ (transformation identique).}$$

Donc si la forme ω^k est fermée, alors $d(p\omega^k) = \omega^k$.

E x e r c i c e. Soient X un champ de vecteurs sur M et ω une k-forme différentielle. Définissons une $(k-1)$-forme différentielle $i_X\omega$ (appelée *produit intérieur* de X par ω) par l'expression

$$(i_X\omega)(\xi_1, \ldots, \xi_{k-1}) = \omega(X, \xi_1, \ldots, \xi_{k-1}).$$

Démontrer la *formule d'homotopie*

$$i_X d + d i_X = L_X,$$

où L_X est l'opérateur de différentiation suivant la direction du champ X.

[L_X opère sur la forme ω au moyen du flot $\{g^t\}$ du champ X ; en formule cela donne :

$$(L_X\omega)(\xi) = \frac{d}{dt}\bigg|_{t=0} \omega(g^t_* \xi).$$

L'opérateur L_X est appelé *dérivation de Lie* ou *dérivation du pêcheur* : un pêcheur assis sur la berge dérive les corps portés par le flot.]

Indication. Désignons par H l'« opérateur d'homotopie » qui à une k-chaîne $\gamma: \sigma \to M$ associe une $(k+1)$-chaîne $H\gamma: (I \times \sigma) \to M$ par la formule $(H\gamma)(t, x) = g^t\gamma(x)$ (où $I = [0, 1]$). Alors

$$g^1\gamma - \gamma = \partial(H\gamma) + H(\partial\gamma).$$

Exercice. Démontrer la formule de différentiation du produit vectoriel sur l'espace euclidien $\mathbf{R}^3$ (ou sur une variété riemannienne) :

$$\operatorname{rot}[\boldsymbol{a}, \boldsymbol{b}] = \{\boldsymbol{a}, \boldsymbol{b}\} + \boldsymbol{a} \operatorname{div} \boldsymbol{b} - \boldsymbol{b} \operatorname{div} \boldsymbol{a}$$

(où $\{\boldsymbol{a}, \boldsymbol{b}\} = L_a \boldsymbol{b}$ est le crochet de Poisson des champs de vecteurs, cf. § 39).

Indication. Si τ est l'élément de volume, alors

$$i_{\operatorname{rot}[\boldsymbol{a}, \boldsymbol{b}]}\tau = di_a i_b \tau, \quad \operatorname{div} \boldsymbol{a} = di_a \tau, \quad \{\boldsymbol{a}, \boldsymbol{b}\} = L_a \boldsymbol{b};$$

en utilisant ces expressions et en se rappelant que $d\tau = 0$, de la formule d'homotopie on déduit aisément la formule de rot $[\boldsymbol{a}, \boldsymbol{b}]$.

G. Appendice 3. Cohomologie et homologie. Toutes les k-formes sur M constituent un espace vectoriel, les k-formes fermées, un sous-espace de cet espace et les différentielles des $(k+1)$-formes, un sous-espace de l'espace des formes fermées. L'espace quotient

$$(\text{formes fermées})/(\text{différentielles}) = H^k(M, \mathbf{R})$$

est appelé *groupe de cohomologie de dimension k* de la variété M. Un élément de ce groupe est une classe de formes fermées différant l'une de l'autre par une différentielle.

Exercice 14. Montrer que $H^1(S^1, \mathbf{R}) = \mathbf{R}$, où S^1 est un cercle.

La dimension de l'espace $H^k(M, \mathbf{R})$ s'appelle *nombre de Betty de dimension k* de la variété M.

Exercice 15. Trouver le nombre de Betty de dimension 1 du tore $T^2 = S^1 \times S^1$.

Le flux d'un fluide (sans sources) à travers la surface de deux sphères concentriques est le même. D'une façon générale, lorsqu'on

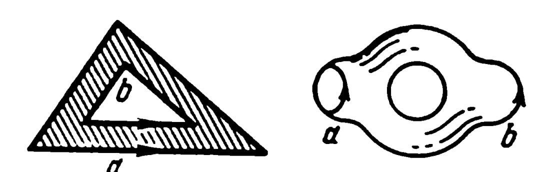


Fig. 165. Cycles homologues.

intègre une forme fermée sur un cycle de dimension k, on peut remplacer ce dernier par un autre sous réserve que leur différence soit

le bord d'une ($k + 1$)-chaîne (fig. 165):

$$\int_a \omega^k = \int_b \omega^k,$$

si $a - b = \partial c_{k+1}$ et $d\omega^k = 0$.

Poincaré a qualifié *d' homologues* deux tels cycles a, b.

Si le groupe des chaînes sur une variété M et les sous-groupes de cycles et de bords (i.e. les cycles homologues à zéro) contenus dans M sont dûment définis *), le groupe quotient

$$(\text{cycles})/(\text{bords}) = H_k(M)$$

s'appelle *groupe d'homologie de dimension k de M*.

Un élément de ce groupe est une classe de cycles homologues.

Le rang de ce groupe est aussi égal au nombre de Betty de dimension k de la variété M (« théorème de De Rama »).

---

*) Il faudrait pour cela réduire notre groupe $\{c_k\}$ en identifiant entre eux les morceaux qui ne diffèrent que par le choix de la paramétrisation f et des polyèdres D. En particulier, on peut considérer que D est toujours le même simplexe ou le même cube. Par ailleurs on pourrait considérer comme nul tout morceau (D, f, Or) de dimension k s'il est dégénéré, i. e. si $f = f_2 \cdot f_1$, où $f_1 : D \longrightarrow D'$ et D' est de dimension inférieure à k.

CHAPITRE 8

VARIÉTÉS SYMPLECTIQUES

Une structure symplectique sur une variété est une 2-forme différentielle fermée régulière sur cette variété. Les espaces des phases des systèmes mécaniques possèdent des structures symplectiques naturelles.

Sur une variété symplectique comme d'ailleurs sur une variété riemannienne il existe un isomorphisme naturel entre les champs de vecteurs et les 1-formes. Un champ de vecteurs sur une variété symplectique correspondant à la différentielle d'une fonction est appelé champ de vecteurs hamiltonien. Un champ de vecteurs sur une variété définit un flot: un groupe à un paramètre de difféomorphismes. Le flot d'un champ de vecteurs hamiltonien sur une variété symplectique conserve la structure symplectique de l'espace des phases.

Les champs de vecteurs sur une variété forment une algèbre de Lie. C'est le cas également des champs de vecteurs hamiltoniens sur une variété symplectique. Les opérations définies sur ces algèbres sont appelées crochets de Poisson.

§ 37. Structure symplectique sur une variété

On définit ici les variétés symplectiques, les champs de vecteurs hamiltoniens sur ces variétés et la structure symplectique standard sur le fibré cotangent.

A. Définition. Soit M une variété différentiable de dimension paire.

Une *structure symplectique* sur M est une 2-forme différentielle fermée et régulière ω^2 sur M:

$$d\omega^2 = 0, \quad \forall \xi \neq 0 \quad \exists \eta : \omega^2(\xi, \eta) \neq 0 \quad (\xi, \eta \in TM_x).$$

Le couple (M, ω^2) s'appelle *variété symplectique*.

Exemple. Considérons un espace vectoriel $\mathbf{R}^{2n}$ muni des coordonnées p_i, q_i et soit $\omega^2 = \Sigma dp_i \wedge dq_i$.

Exercice. Vérifier que $(\mathbf{R}^{2n}, \omega^2)$ est une variété symplectique. Pour $n = 1$ le couple $(\mathbf{R}^2, \omega^2)$ est le couple (plan, aire).

L'exemple qui va suivre explique comment les variétés symplectiques apparaissent en dynamique. En dehors du fibré tangent d'une variété différentiable il est utile de considérer le fibré cotangent dual du fibré tangent.

B. Le fibré cotangent et sa structure symplectique. Soit V une variété différentiable de dimension n. On appelle *vecteur cotangent à V en un point $x \in V$* 1-forme sur l'espace tangent à V en x. L'ensemble de tous les vecteurs cotangents à V en x est un espace vectoriel de dimension n dual de l'espace tangent TV_x. Cet espace est noté T^*V_x et s'appelle *espace cotangent* à la variété V en x.

L'union de tous les espaces cotangents à la variété V en tous ses points s'appelle *fibré cotangent de V* et se note T^*V. L'espace T^*V possède une structure naturelle de variété différentiable de dimension $2n$. Un point de T^*V est une 1-forme sur l'espace tangent à V en l'un de ses points. Si $\mathbf{q}$ est une collection de n coordonnées locales d'un point de V, alors une telle forme est définie par la donnée de ses n composantes $\mathbf{p}$. Pris ensemble, les $2n$ nombres $\mathbf{p}$ et $\mathbf{q}$ forment une collection de coordonnées locales d'un point de T^*V.

Il existe une application projective $f: T^*V \to V$ (associant à toute 1-forme sur TV_x un point x de V). L'application f est une application différentiable *sur*. L'antécédent d'un point $x \in V$ par l'application f est l'espace cotangent T^*V_x.

Théorème. *Le fibré cotangent T^*V possède une structure symplectique naturelle. Dans les coordonnées locales décrites plus haut, cette structure est donnée par la formule*

$$\omega^2 = d\mathbf{p} \wedge d\mathbf{q} = dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n.$$

D é m o n s t r a t i o n. Pour commencer on va définir une 1-forme remarquable sur T^*V. Soit $\xi \in T(T^*V)_p$ un vecteur *tangent au fibré cotangent en un point $p \in T^*V_x$* (fig. 166). La dérivée $f_*: T(T^*V) \to TV$ de l'application projective $f: T^*V \to V$ envoie le vecteur ξ dans le vecteur $f_*\xi$ tangent en x à V. Définissons une 1-forme ω^1 sur T^*V par la relation $\omega^1(\xi) = p(f_*\xi)$. Dans les coordonnées locales introduites plus haut cette forme s'écrit $\omega^1 = \mathbf{p}\, d\mathbf{q}$. En vertu de l'exemple du point A la 2-forme fermée $\omega^2 = d\omega^1$ est régulière.

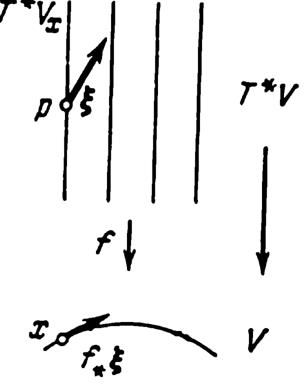


Fig. 166. 1-forme $\mathbf{p}\, d\mathbf{q}$ sur le fibré cotangent.

R e m a r q u e. Considérons un système mécanique lagrangien de variété de configuration V et de lagrangien L. Il est aisé de voir que la « vitesse généralisée » lagrangienne $\dot{\mathbf{q}}$ est un vecteur tangent à la variété de configuration V,

et l'« impulsion généralisée » $p = \partial L/\partial \dot{q}$ un vecteur cotangent. Donc l'espace des phases $\{(p, q)\}$ du problème lagrangien est le fibré cotangent à la variété de configuration. Par conséquent le théorème précédent montre que l'espace des phases d'un problème de mécanique possède une structure naturelle de variété symplectique.

E x e r c i c e. Montrer que la transformation de Legendre ne dépend pas du système de coordonnées: elle associe à une fonction $L: TV \to \mathbf{R}$ sur le fibré tangent une fonction $H: T^*V \to \mathbf{R}$ sur le cotangent.

C. Champs de vecteurs hamiltoniens. Une structure riemannienne sur une variété définit un isomorphisme entre l'espace des vecteurs tangents et celui des 1-formes. Une structure symplectique définit un isomorphisme analogue.

D é f i n i t i o n. Associons à un vecteur ξ tangent en x à une variété symplectique (M, ω^2) une 1-forme ω_ξ^1 sur TM_x moyennant la formule

$$\omega_\xi^1(\eta) = \omega^2(\eta, \xi) \quad \forall \eta \in TM_x.$$

E x e r c i c e. Montrer que la correspondance $\xi \to \omega_\xi^1$ est un isomorphisme de l'espace des vecteurs sur celui des 1-formes.

E x e m p l e. Dans $\mathbf{R}^{2n} = \{(p, q)\}$ on identifiera les vecteurs et les 1-formes conformément à la structure euclidienne $(x, x) = p^2+q^2$. Alors la correspondance $\xi \to \omega_\xi^1$ définira une transformation $\mathbf{R}^{2n} \to \mathbf{R}^{2n}$.

E x e r c i c e. Calculer la matrice de cette transformation dans la base p, q.

Réponse. $\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$.

Désignons par I l'isomorphisme construit plus haut: $I: T^*M_x \to TM_x$.

Soit maintenant H une fonction sur la variété symplectique M. Alors dH est une 1-forme différentielle sur M à laquelle en chaque point est associé un vecteur tangent à M. Nous obtenons ainsi un champ de vecteurs $I\,dH$ sur M.

D é f i n i t i o n. Le champ de vecteurs $I\,dH$ s'appelle *champ de vecteurs hamiltonien, H est la fonction de Hamilton ou hamiltonien.*

E x e m p l e. Si $M = \mathbf{R}^{2n} = \{(p, q)\}$, alors on obtient le champ de vitesses de phase des équations canoniques de Hamilton

$$\dot{x} = I\,dH(x) \Leftrightarrow \dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}.$$

§ 38. Flots hamiltoniens et leurs invariants intégraux

Le théorème de Liouville affirme que le flot conserve le volume. Poincaré a trouvé de nombreuses formes différentielles invariantes par le flot hamiltonien.

A. Les flots hamiltoniens conservent la structure symplectique. Soit (M, ω^2) une variété symplectique, $H: M \to \mathbf{R}$ une fonction.

Supposons que le champ de vecteurs hamiltonien $I\,dH$ correspondant à H définisse un groupe à un paramètre de difféomorphismes $g^t: M \to M$,

$$\frac{d}{dt}\Big|_{t=0} g^t x = I\,dH(x).$$

Le groupe g^t est appelé *flot hamiltonien de fonction de Hamilton H.*

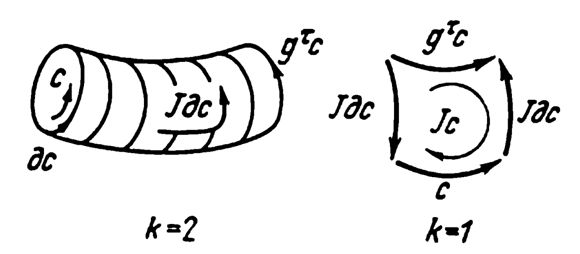


Fig. 167. Trace d'une chaîne par homotopie.

Théorème. *Le flot hamiltonien conserve la structure symplectique:*

$$(g^t)^* \omega^2 = \omega^2.$$

Pour $n = 1$, $M = \mathbf{R}^2$ ce théorème dit que le flot g^t conserve l'aire (théorème de Liouville).

Pour les besoins de la démonstration introduisons les notations suivantes (fig. 167).

Soit M une variété quelconque, c une k-chaîne sur M, $g^t: M \to M$ une famille d'applications différentiables à un paramètre. Construisons une $(k+1)$-chaîne Jc sur M que nous appellerons *trace de la chaîne c par l'homotopie* g^t, $0 \leqslant t \leqslant \tau$.

Soit (D, f, Or) un morceau de chaîne c. Dans la chaîne Jc il lui correspondra le morceau (D', f', Or'), où $D' = I \times D$ est le produit cartésien de l'intervalle $0 \leqslant t \leqslant \tau$ par D, l'application $f': D' \to M$ est exprimée en fonction de $f: D \to M$ au moyen de la formule $f'(t, x) = g^t f(x)$ et l'orientation Or$'$ de l'espace $\mathbf{R}^{k+1}$ contenant D' est définie par le repère $e_0, e_1, \ldots, e_k$, où e_0 est le vecteur unitaire de l'axe t et $e_1, \ldots, e_k$ le repère orientant D.

On pourrait dire aussi que Jc est la chaîne *balayée* par c par l'homotopie g^t, $0 \leqslant t \leqslant \tau$. Le bord de la chaîne Jc est composé des positions initiale et finale de c et d'une « surface latérale » balayée par le bord de c.

On vérifie sans difficultés qu'avec les orientations choisies

$$\partial (Jc_k) = g^\tau c_k - c_k - J\partial c_k.$$

Lemme. *Soit γ une 1-chaîne sur une variété symplectique (M, ω^2) et soit g^t le flot sur M de hamiltonien H. Alors*

$$\frac{d}{d\tau}\int_{J\gamma} \omega^2 = \int_{g^\tau \gamma} dH.$$

D é m o n s t r a t i o n. Il suffit de considérer une chaîne γ constituée d'un seul morceau $f: [0, 1] \to M$.

Introduisons les notations

$$f'(s, t) = g^t f(s), \quad \xi = \partial f'/\partial s, \quad \eta = \partial f'/\partial t \in TM_{f'(s,\, t)}.$$

Par définition de l'intégrale

$$\int_{J\gamma} \omega^2 = \int_0^1 \int_0^\tau \omega^2(\eta, \xi)\, dt\, ds.$$

Or par définition du flot, η est un vecteur du champ de hamiltonien H (appliqué au point $f'(s, t)$). Par définition du champ hamiltonien $\omega^2(\eta, \xi) = dH(\xi)$. Donc

$$\int_{J\gamma} \omega^2 = \int_0^\tau \Big(\int_{g^t\gamma} dH \Big)\, d\tau.$$

Le lemme est démontré.

Corollaire. *Si la chaîne γ est fermée* ($\partial\gamma = 0$), *alors* $\int_{J\gamma} \omega^2 = 0$.

En effet $\int_\gamma dH = \int_{\partial\gamma} H = 0$.

D é m o n s t r a t i o n d u t h é o r è m e. Soit une 2-chaîne quelconque c. On a

$$0 \overset{1}{=} \int_{Jc} d\omega^2 \overset{2}{=} \int_{\partial Jc} \omega^2 \overset{3}{=} \int_{g^\tau c} - \int_c - \int_{J\partial c} \omega^2 \overset{4}{=} \int_{g^\tau c} \omega^2 - \int_c \omega^2$$

(l'égalité 1 a lieu du fait que ω^2 est fermée, 2 est la formule de Stokes, 3 la formule (1), 4 est vraie en vertu du corollaire précédent, $\gamma = \partial c$). Donc les intégrales de la forme ω^2 sur une chaîne quelconque c et sur son image $g^\tau c$ sont égales, c.q.f.d.

E x e r c i c e. Tout groupe à un paramètre de difféomorphismes de M^{2n} conservant la structure symplectique est-il un flot hamiltonien?

I n d i c a t i o n. Cf. § 40.

B. Invariants intégraux. Soit $g: M \to M$ une application différentiable.

D é f i n i t i o n. Une k-forme différentielle ω est par définition un *invariant intégral* par l'application g si les intégrales de ω sur une k-chaîne quelconque c et sur son image par l'application g sont égales:

$$\int_{gc} \omega = \int_c \omega.$$

Exemple. Si $M = \mathbf{R}^2$ et $\omega^2 = dp \wedge dq$ est l'élément de surface, alors ω^2 est invariant intégral par toute application g de jacobien 1.

Exercice. Montrer qu'une forme ω^k est invariant intégral par une application g si et seulement si $g^*\omega^k = \omega^k$.

Exercice. Montrer que si les formes ω^k et ω^l sont invariants intégraux par une application g, alors la forme $\omega^k \wedge \omega^l$ l'est également.

On peut donner une autre formulation au théorème du point A.

Théorème. *Une forme ω^2 définissant une structure symplectique est invariant intégral par le flot hamiltonien.*

Voyons maintenant les puissances extérieures de ω^2:

$$(\omega^2)^2 = \omega^2 \wedge \omega^2, \quad (\omega^2)^3 = \omega^2 \wedge \omega^2 \wedge \omega^2, \ \ldots .$$

Corollaire. *Chacune des formes* $(\omega^2)^2$, $(\omega^2)^3$, ... *est invariant intégral par le flot hamiltonien.*

Exercice. Soit une variété symplectique (M^{2n}, ω^2) de dimension $2n$. Montrer que $(\omega^2)^k = 0$ pour $k > n$ et que $(\omega^2)^n$ est une $2n$-forme régulière sur M^{2n}.

Définissons l'élément de volume sur M à l'aide de $(\omega^2)^n$. Le flot hamiltonien conserve alors le volume et le corollaire précédent entraîne théorème de Liouville.

Exemple. Considérons l'espace symplectique arithmétique $M = \mathbf{R}^{2n} = \{(p, q)\}$, $\omega^2 = dp \wedge dq = \sum dp_i \wedge dq_i$. Dans ce cas la forme $(\omega^2)^k$ est proportionnelle à la forme

$$\omega^{2k} = \sum_{i_1 < \ldots < i_k} dp_{i_1} \wedge \ldots \wedge dp_{i_k} \wedge dq_{i_1} \wedge \ldots \wedge dq_{i_k}.$$

L'intégrale de la forme ω^{2k} sur une chaîne est égale à la somme des volumes orientés des projections de cette chaîne sur les plans arithmétiques $(p_{i_1}, \ldots, p_{i_k}, q_{i_1} \ldots, q_{i_k})$.

Une application $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ est dite *canonique* si elle admet ω^2 pour invariant intégral. Chacune des formes $\omega^4, \omega^6, \ldots, \omega^{2n}$ est invariant intégral par toute application canonique. Donc la *somme des aires orientées des projections des chaînes sur les plans arithmétiques* $(p_{i_1}, \ldots, p_{i_k}, q_{i_1}, \ldots, q_{i_k})$, $1 \leqslant k \leqslant n$, *est invariante par toute application canonique*. En particulier, *les applications canoniques conservent le volume.*

Le flot hamiltonien défini par les équations $\dot{p} = -\frac{\partial H}{\partial q}$, $\dot{q} = \frac{\partial H}{\partial p}$ est composé d'applications canoniques g^t.

Les invariants intégraux que nous avons vus s'appellent également *invariants absolus.*

D é f i n i t i o n. Une k-forme différentielle ω est *invariant intégral relatif* par une application $g: M \to M$ si

$$\int_{gc} \omega = \int_{c} \omega$$

sur toute k-chaîne c *fermée*

Théorème. *Soit ω un invariant intégral relatif par une application g; alors $d\omega$ est invariant intégral absolu par g.*

D é m o n s t r a t i o n. Soit c une $(k+1)$-chaîne. Il vient

$$\int_{c} d\omega \overset{1}{=} \int_{\partial c} \omega \overset{2}{=} \int_{g\partial c} \omega \overset{3}{=} \int_{\partial gc} \omega \overset{4}{=} \int_{gc} d\omega$$

(1 et 4 découlent de la formule de Stokes, 2 de la définition de l'invariant relatif, 3 de la définition du bord).

E x e m p l e. *Toute application canonique* $g: R^{2n} \to R^{n2}$ *admet pour invariant intégral relatif la 1-forme*,

$$\omega^1 = p\, dq - \sum_{i=1}^{n} p_i\, dq_i.$$

En effet, toute chaîne fermée c dans R^{2n} est bord d'une certaine chaîne σ et l'on a

$$\int_{gc} \omega^1 \overset{1}{=} \int_{g\partial\sigma} \omega^1 \overset{2}{=} \int_{\partial g\sigma} \omega^1 \overset{3}{=} \int_{g\sigma} d\omega^1 \overset{4}{=} \int_{\sigma} d\omega^1 \overset{5}{=} \int_{\partial\sigma} \omega^1 \overset{6}{=} \int_{c} \omega^1$$

(1 et 6 découlent de la définition de σ, 2 de la définition de ∂, 3 et 5 de la formule de Stokes, 4 de la canonicité de g et de $d\omega^1 = d\,(p\, dq) = dp \wedge dq = \omega^2$).

E x e r c i c e. Soit $d\omega^k$ un invariant intégral absolu par une application $g: M \to M$. En résulte-t-il que ω^k est invariant intégral relatif?

Réponse. Non, s'il existe des k-chaînes fermées dans M qui ne sont pas bords.

C. Loi de conservation de l'énergie.

Théorème. *Le flot hamiltonien de fonction de Hamilton H admet H pour intégrale première.*

D é m o n s t r a t i o n. La dérivée de H suivant la direction d'un vecteur η est égale à la valeur de dH sur le vecteur η. Par définition du champ hamiltonien $\eta = I\, dH$ on trouve

$$dH\,(\eta) = \omega^2\,(\eta, I dH) = \omega^2\,(\eta, \eta) = 0.$$

E x e r c i c e. Montrer que la 1-forme dH est invariant intégral par le flot de hamiltonien H.

§ 39. Algèble de Lie des champs de vecteurs

A tout couple de champs de vecteurs sur une variété est associé un nouveau champ de vecteurs appelé crochet de Poisson de ces champs. Le crochet de Poisson transforme l'espace vectoriel des champs de vecteurs infiniment différentiables sur une variété en algèbre de Lie.

A. Algèbre de Lie. Un exemple d'algèbre de Lie est un espace vectoriel euclidien orienté muni de la multiplication vectorielle. Le produit vectoriel est bilinéaire, antisymétrique et vérifie l'identité de Jacobi

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0.$$

D é f i n i t i o n. On appelle *algèbre de Lie* un espace vectoriel L muni d'une opération bilinéaire antisymétrique $L \times L \to L$ vérifiant l'identité de Jacobi.

Cette opération est notée par un crochet et s'appelle *commutateur*.

E x e r c i c e. Montrer que l'ensemble des matrices $n \times n$ est une algèbre de Lie si on définit le commutateur comme $[A, B] = AB - BA$.

B. Champs de vecteurs et opérateurs différentiels. Soient M une variété différentiable, $\mathbf{A}$ un champ de vecteurs différentiable sur M: en chaque point $x \in M$ est donné un vecteur $\mathbf{A}(x) \in TM_x$ tangent à M en x. A tout tel champ de vecteurs sont liés:

1. Un *groupe à un paramètre* *) *de difféomorphismes* ou le *flot* $A^t: M \to M$ pour lequel $\mathbf{A}$ est champ de vitesses (fig. 168):

$$\frac{d}{dt}\Big|_{t=0} A^t(x) = \mathbf{A}(x).$$

2. *L'opérateur différentiel L_A d'ordre* 1. Il s'agit de la différentiation des fonctions suivant la direction du champ $\mathbf{A}$: pour toute

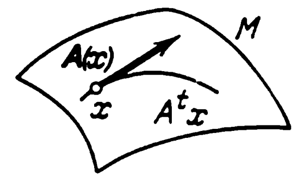


Fig. 168. Groupe de difféomorphismes défini par un champ de vecteurs.

fonction $\varphi: M \to \mathbf{R}$ la dérivée suivant la direction du champ $\mathbf{A}$ est une fonction $L_A\varphi$ qui vaut au point x

$$(L_A\varphi)(x) = \frac{d}{dt}\Big|_{t=0} \varphi(A^t(x)).$$

*) D'après les théorèmes d'existence, d'unicité et de différentiabilité de la théorie des équations différentielles, le groupe A^t est défini si la variété M est compacte. Dans le cas général, les applications A^t sont définies seulement dans le voisinage de x et seulement pour les petits t; ceci est suffisant pour les constructions ultérieures.

E x e r c i c e. Montrer que l'opérateur L_A est linéaire:

$$L_A(\lambda_1\varphi_1 + \lambda_2\varphi_2) = \lambda_1 L_A\varphi_1 + \lambda_2 L_A\varphi_2 \ (\lambda_1, \lambda_2 \in \mathbf{R}).$$

Démontrer la formule de Leibniz $L_A(\varphi_1\varphi_2) = \varphi_1 L_A\varphi_2 + \varphi_2 L_A\varphi_1$.

E x e m p l e. Soit $(x_1, \ldots, x_n)$ un système de coordonnées local sur M. Dans ce système le vecteur $\boldsymbol{A}(x)$ a pour composantes $(A_1(x), \ldots, A_n(x))$; le flot A^t est défini par le système d'équations différentielles

$$\dot{x}_1 = A_1(x), \ldots, \dot{x}_n = A_n(x)$$

et donc la dérivée de $\varphi = \varphi(x_1, \ldots, x_n)$ suivant la direction de $\boldsymbol{A}$ s'écrit

$$L_A\varphi = A_1\frac{\partial\varphi}{\partial x^1} + \ldots + A_n\frac{\partial\varphi}{\partial x_n},$$

Autrement dit, dans les coordonnées $(x_1, \ldots, x_n)$ l'opérateur L_A s'écrit

$$L_A = A_1\frac{\partial}{\partial x_1} + \ldots + A_n\frac{\partial x}{\partial x_n},$$

ceci n'est autre que la forme générale de l'opérateur différentiel linéaire du premier ordre dans un espace arithmétique.

E x e r c i c e. Montrer que la correspondance entre les champs de vecteurs $\boldsymbol{A}$, les flots A^t et les différentiations L_A est bijective.

C. Crochet de Poisson de champs de vecteurs.

Soient donnés sur une variété M deux champs de vecteurs $\boldsymbol{A}$ et $\boldsymbol{B}$.

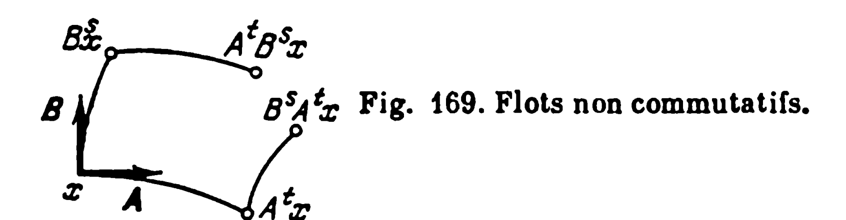


Fig. 169. Flots non commutatifs.

Les flots A^t et B^s correspondants, d'une façon générale, ne commutent pas: $A^tB^s \neq B^sA^t$ (fig. 169).

E x e r c i c e. Citer un exemple.
S o l u t i o n. Les champs $\boldsymbol{A} = e_1$, $\boldsymbol{B} = x_1e_2$ sur le plan (x_1, x_2).

Pour définir le degré de non-commutativité des flots A^t et B^s considérons les points A^tB^sx et B^sA^tx. Pour évaluer la différence qui existe entre ces points comparons les valeurs qu'y prend une fonction différentiable quelconque φ définie sur M. La différence

$$\Delta(t,s;x) = \varphi(A^tB^sx) - \varphi(B^sA^tx)$$

est visiblement une fonction différentiable qui s'annule pour $s = 0$ et $t = 0$. Donc le premier terme de la série de Taylor de Δ non nul en 0 contient le produit st, quant aux autres termes du second degré ils disparaissent. Calculons ce terme principal bilinéaire de Δ en 0.

Lemme 1. *En 0, la dérivée mixte de Δ par rapport à s et t est égale au commutateur de différentiation suivant les directions de A et B:*

$$\frac{\partial^2}{\partial s\,\partial t}\bigg|_{s=t=0} \varphi(A^tB^sx)-\varphi(B^sA^tx)=(L_BL_A\varphi-L_AL_B\varphi)(x).$$

D é m o n s t r a t i o n. Par définition de L_A

$$\frac{\partial}{\partial t}\bigg|_{t=0} \varphi(A^tB^sx)=(L_A\varphi)(B^sx).$$

Si l'on désigne la fonction $L_A\varphi$ par ψ, alors la définition de L_B entraîne

$$\frac{\partial}{\partial s}\bigg|_{s=0} \psi(B^sx)=(L_B\psi)(x).$$

Donc

$$\frac{\partial^2}{\partial s\,\partial t}\bigg|_{s=t=0} \varphi(A^tB^sx)=(L_BL_A\varphi)(x),$$

c.q.f.d.

Voyons maintenant le commutateur de différentiation $L_BL_A - L_AL_B$ engendré. A première vue on dirait que c'est un opérateur différentiel du *second* ordre.

Lemme 2. *L'opérateur $L_BL_A-L_AL_B$ est un opérateur différentiel linéaire du premier ordre.*

D é m o n s t r a t i o n. Soient $(A_1, \ldots, A_n)$ et $(B_1, \ldots, B_n)$ les composantes des champs A et B dans un système de coordonnées local $(x_1, \ldots, x_n)$ sur M. On a

$$L_BL_A\varphi=\sum_{i=1}^{n} B_i\frac{\partial}{\partial x_i}\sum_{j=1}^{n} A_j\frac{\partial}{\partial x_j}\varphi=$$

$$=\sum_{i,j=1}^{n} B_i\frac{\partial A_j}{\partial x_i}\frac{\partial}{\partial x_j}\varphi+\sum_{i,j=1}^{n} B_iA_j\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}.$$

Si l'on retranche $L_AL_B\,\varphi$, le terme qui contient les dérivées secondes de φ disparaît et l'on obtient

$$(L_BL_A-L_AL_B)\varphi=\sum_{i,j=1}^{n}\left(B_i\frac{\partial A_j}{\partial x_i}-A_i\frac{\partial B_j}{\partial x_i}\right)\frac{\partial\varphi}{\partial x_j}.$$

Le lemme est démontré.

D'autre part, vu que tout opérateur différentiel linéaire du premier ordre est défini par un champ de vecteurs, notre opérateur $L_BL_A-L_AL_B$ correspond également à un certain champ de vecteurs C.

D é f i n i t i o n. On appelle *crochet de Poisson* ou *commutateur* *) de deux champs de vecteurs A et B sur une variété M le champ de

*) De nombreux auteurs préconisent un autre symbole. Le nôtre est conforme à celui du commutateur de la théorie des groupes de Lie (voir point F).

vecteurs C tel que

$$L_C = L_B L_A - L_A L_B.$$

Le crochet de Poisson de deux champs de vecteurs est noté

$$C = [A, B].$$

Exercice. Supposons que deux champs A et B sont donnés dans les coordonnées x_i par leurs composantes (A_i, B_i). Trouver les composantes du crochet de Poisson.

Solution. En démontrant le lemme 2 on a établi la formule

$$[A, B]_j = \sum_{i=1}^{n} B_i \frac{\partial A_j}{\partial x_i} - A_i \frac{\partial B_j}{\partial x_i}.$$

Exercice. Soient A_1 et A_2 les champs de vecteurs des vitesses linéaires d'un solide en rotation respectivement avec les vitesses angulaires ω_1 et ω_2 autour d'un point O. Calculer le crochet de Poisson $[A_1, A_2]$.

D. Identité de Jacobi.

Théorème. *Le crochet de Poisson transforme l'espace vectoriel des champs de vecteurs sur une variété M en algèbre de Lie.*

Démonstration. La linéarité et l'antisymétrie du crochet de Poisson sont évidentes. Prouvons l'identité de Jacobi. Par définition du crochet de Poisson on a

$$L_{[[A, B], C]} = L_C L_{[A, B]} - L_{[A, B]} L_C =$$
$$= L_C L_B L_A - L_C L_A L_B + L_A L_B L_C - L_B L_C L_A.$$

La somme $L_{[[A, B], C]} + L_{[[B, C], A]} + L_{[[C, A], B]}$ comprendra en tout 12 termes. Chaque terme y figurera deux fois avec des signes opposés. Le théorème est démontré.

E. Condition de commutativité des flots. Soient A, B des champs de vecteurs sur une variété M.

Théorème. *Deux flots A^t, B^s commutent si et seulement si le crochet de Poisson $[A, B]$ des champs de vecteurs correspondants est nul.*

Démonstration. Si $A^t B^s \equiv B^s A^t$, alors en vertu du lemme 1 on a $[A, B] = 0$. Si $[A, B] = 0$, alors en vertu du lemme 1 pour toute fonction φ en un point quelconque x on a

$$\varphi (A^t B^s x) - \varphi (B^s A^t x) = o (s^2 + t^2), \quad s \to 0, \ t \to 0.$$

Nous allons montrer que ceci entraîne $\varphi (A^t B^s x) = \varphi (B^s A^t x)$ pour s et t suffisamment petits.

En appliquant cette relation aux coordonnées locales $(\varphi = x_1, \ldots, \varphi = x_n)$ on obtient $A^t B^s \equiv B^s A^t$.

Considérons le rectangle $0 \leqslant t \leqslant t_0$, $0 \leqslant s \leqslant s_0$ (fig. 170) sur le plan (t, s). A chaque chemin joignant $(0, 0)$ et (t_0, s_0) et composé d'un nombre fini de segments parallèles aux axes de coordonnées associons un produit de transformations A^t et B^s. A tout segment $t_1 \leqslant t \leqslant t_2$ associons $A^{t_2 - t_1}$ et à $s_1 \leqslant s \leqslant s_2$,

$B^{s_2-s_1}$; nous appliquerons les transformations dans le sens qui va de (0, 0) à (t_0, s_0).

Ainsi aux côtés $(0 \leqslant t \leqslant t_0, s = 0)$ et $(t = t_0, 0 \leqslant s \leqslant s_0)$ est associé le produit $B^{s_0}A^{t_0}$, et aux côtés $(t = 0, 0 \leqslant s \leqslant s_0)$ et $(s = s_0, 0 \leqslant t \leqslant t_0)$, le produit $A^{t_0}B^{s_0}$.

En outre à tout tel chemin sur le plan (t, s) on fait correspondre un chemin sur M issu de x et composé des trajectoires des flots A^t et B^s (fig. 171). Si à un chemin sur le plan (t, s) est associée la transformation $A^{t_1}B^{s_1} \ldots A^{t_n}B^{s_n}$, alors sur M le chemin correspondant s'achève au point $A^{t_1}B^{s_1} \ldots A^{t_n}B^{s_n}x$.

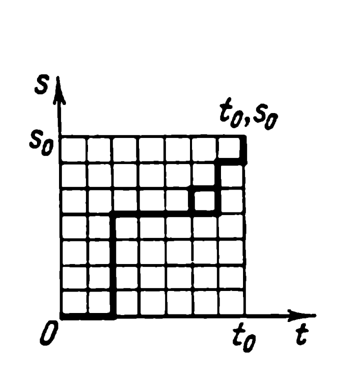


Fig. 170. Sur la commutativité des flost.

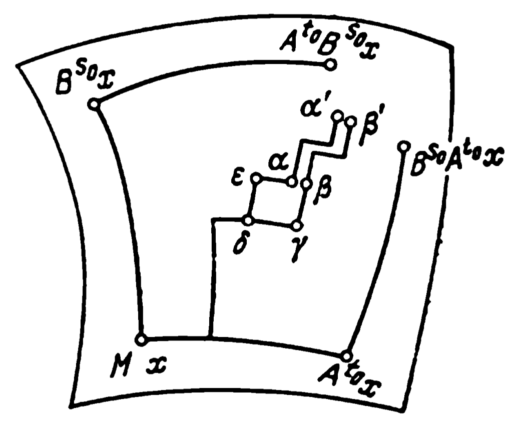


Fig. 171. Quadrilatère curviligne βγδεα.

L'on se propose de démontrer que tous ces chemins finissent en réalité au point $A^{t_0}B^{s_0}x = B^{s_0}A^{t_0}x$.

Partageons les segments $(0 \leqslant t \leqslant t_0)$ et $(0 \leqslant s \leqslant s_0)$ en N parties égales, de sorte que tout le rectangle se trouve divisé en N^2 petits rectangles. Pour passer des côtés $(0, 0) — (0, t_0) — (s_0, t_0)$ aux côtés $(0, 0) — (s_0, 0) — (s_0, t_0)$

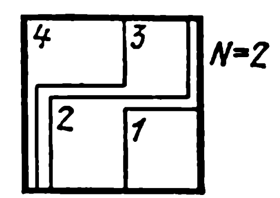


Fig. 172. Passage d'un couple de côtés à un autre.

on peut effectuer N^2 pas au cours de chacun desquels deux côtés adjacents d'un petit rectangle sont remplacés par les deux autres (fig. 172).

Sur la variété M à ce petit rectangle correspond d'une façon générale un quadrilatère curviligne non fermé βγδεα (fig. 171). Considérons la distance *) entre ses sommets α et β correspondant aux plus grandes valeurs de t et s. Nous avons vu plus haut que $\rho(\alpha, \beta) \leqslant C_1N^{-3}$ (où la constante $C_1 > 0$ ne dépend pas de N). En utilisant le théorème de différentiabilité des solutions des équations différentielles par rapport aux conditions initiales on déduit sans peine la majoration de la distance des extrémités α' et β' des chemins $x\delta\gamma\beta\beta'$ et $x\delta\varepsilon\alpha\alpha'$ sur la variété M: $\rho(\alpha', \beta') < C_2N^{-3}$, où la constante $C_2 > 0$ ne dépend pas de N non plus. Or pour passer de $B^{s_0}A^{t_0}x$ à $A^{t_0}B^{s_0}x$ nous avons partagé le chemin en N^2 pas. Donc $\rho(A^{t_0}B^{s_0}x, B^{s_0}A^{t_0}x) \leqslant N^2C_2N^{-3}\ \forall N$. $A^{t_0}B^{s_0}x = B^{s_0}A^{t_0}x$.

*) Dans une métrique riemannienne quelconque sur M.

F. Appendice. Algèbre de Lie d'un groupe de Lie. Un groupe G est par définition un *groupe de Lie* si G est une variété différentiable et si les opérations (produit et inversion) sont des applications différentiables $C \times G \to G$, $G \to G$.

L'espace TG_e tangent en l'unité au groupe de Lie G possède une structure naturelle d'algèbre de Lie définie de la manière suivante.

A tout vecteur tangent $\boldsymbol{A} \in TG_e$ correspond un sous-groupe $A^t \in \in G$ de vecteur vitesse $\boldsymbol{A} = \frac{d}{dt}\Big|_{t=0} A^t$.

La non-commutativité de deux sous-groupes A^t et B^s se mesure par le produit $A^tB^sA^{-t}B^{-s}$. Il s'avère qu'il existe un sous-groupe C^r et un seul tel que

$$\rho\,(A^tB^sA^{-t}B^{-s},\ C^{st}) = o\,(s^2 + t^2) \text{ pour } s,\ t \to 0.$$

Le vecteur correspondant $\boldsymbol{C} = \frac{d}{dr}\Big|_{r=0} C^r$ s'appelle *commutateur de Lie* $\boldsymbol{C} = [\boldsymbol{A}, \boldsymbol{B}]$ des vecteurs $\boldsymbol{A}$ et $\boldsymbol{B}$.

On vérifie que l'opération de commutation dont on vient de munir l'espace tangent TG_e transforme cet espace en algèbre de Lie (i.e. cette opération est bilinéaire, antisymétrique et vérifie l'identité de Jacobi). Cette algèbre s'appelle *algèbre de Lie d'un groupe de Lie G*.

E x e r c i c e. Etablir l'expression de la commutation d'une algèbre de Lie du groupe SO (3) des rotations de l'espace euclidien $\mathbf{R}^3$.

Le lemme 1 dit que le *crochet de Poisson de champs de vecteurs peut être défini comme le commutateur de Lie du « groupe de Lie de dimension infinie » de tous les difféomorphismes de la variété* *) *M*.

D'un autre côté le commutateur de Lie peut être défini à l'aide du crochet de Poisson de champs de vecteurs sur un groupe de Lie G.

Soit $g \in G$. On appelle *translation à droite* R_g une application R_g: $G \to G$, $R_gh = hg$. La différentielle de R_g au point g envoie TG_e dans TG_g. Donc à tout vecteur $\boldsymbol{A} \in TG_e$ est associé tout un champ de vecteurs sur le groupe : ce champ est composé de toutes les translations à droite $(R_g)_*\boldsymbol{A}$ et s'appelle *champ invariant à droite*. De toute évidence un champ invariant à droite sur un groupe est univoquement défini par sa valeur à l'unité.

E x e r c i c e. Montrer que le *crochet de Poisson de champs de vecteurs invariants à droite sur un groupe de Lie G est un champ de vecteurs invariant à droite et que sa valeur à l'unité du groupe est égale au commutateur de Lie des valeurs des champs initiaux à l'unité.*

---

*) Le signe du crochet de Poisson de champs de vecteurs a été choisi en conséquence.

§ 40. Algèbre de Lie des fonctions de Hamilton

Les champs de vecteurs hamiltoniens sur une variété symplectique forment une sous-algèbre de l'algèbre de Lie de tous les champs. Les fonctions de Hamilton constituent également une algèbre de Lie: l'opération de cette algèbre s'appelle crochet de Poisson des fonctions. Les intégrales premières d'un flot hamiltonien sont une sous-algèbre de l'algèbre de Lie des fonctions de Hamilton.

A. Crochet de Poisson de deux fonctions. Soit (M, ω^2) une variété symplectique. A une fonction $H: M \to \mathbf{R}$ définie sur cette variété correspond un groupe à un paramètre $g_H^t: M \to M$ de transformations canoniques de M, notamment le flot de hamiltonien H.

Soit $F: M \to \mathbf{R}$ une autre fonction sur la variété M.

D é f i n i t i o n. On appelle *crochet de Poisson* (F, H) *des fonctions* F et H définies sur la variété symplectique (M, ω^2) la dérivée de la fonction F suivant la direction du flot de hamiltonien H

$$(F, H)(x) = \frac{d}{dt}\Big|_{t=0} F(g_H^t(x)).$$

Donc le crochet de Poisson de deux fonctions définies sur M est de nouveau une fonction sur M.

Corollaire 1. *Une condition nécessaire et suffisante pour qu'une fonction F soit intégrale première du flot de hamiltonien H est que le crochet de Poisson de F et H soit identiquement nul*: $(H, F) \equiv 0$.

On peut donner une autre forme à la définition du crochet de Poisson si l'on se sert de l'isomorphisme I entre les 1-formes et les champs de vecteurs sur la variété symplectique (M, ω^2). Cet isomorphisme est défini par la relation (cf. § 37)

$$\omega^2(\eta, I\omega^1) = \omega^1(\eta).$$

Le vecteur vitesse du flot g_H^t est $I\, dH$. Il s'ensuit le

Corollaire 2. *Le crochet de Poisson des fonctions F et H est égal à la valeur de la 1-forme dF sur le vecteur vitesse $I\, dH$ du flot de hamiltonien H*:

$$(F, H) = dF\,(I\, dH).$$

En appliquant une seconde fois la formule précédente on obtient le

Corollaire 3. *Le crochet de Poisson des fonctions F et H est égal au «produit scalaire gauche»* *) *des vecteurs vitesses des flots de hamiltoniens H et F*:

$$(F, H) = \omega^2(I\, dH, I\, dF).$$

*) ou « skew product ». — *Note du traducteur.*

Alors est évident le

Corollaire 4. *Le crochet de Poisson des fonctions F et H est une fonction antisymétrique bilinéaire de F et H:* $(F, H) = -(H, F)$, $(H, \lambda_1 F_1 + \lambda_2 F_2) = \lambda_1 (H, F_1) + \lambda_2 (H, F_2)$ $(\lambda_i \in \mathbf{R})$.

Tout évidents qu'ils soient, les raisonnements précédents aboutissent à des conclusions non triviales et notamment à la généralisation suivante du théorème de E. Noether.

Théorème. *Si une fonction de Hamilton H donnée sur une variété symplectique (M, ω^2) est invariante par le groupe à un paramètre de transformations canoniques engendré par un hamiltonien F, alors F est intégrale première du système de hamiltonien H.*

En effet, par hypothèse H est intégrale première du flot g_F^t, d'où $(F, H) = 0$ (corollaire 1). Donc $(H, F) = 0$ (corollaire 4) et F est intégrale première (corollaire 1), c.q.f.d.

Exercice 1. Calculer le crochet de Poisson des fonctions F et H dans l'espace arithmétique canonique $\mathbf{R}^{2n} = \{(p, q)\}$, $\omega^2(\xi, \eta) = [\xi, \eta] = (I\xi, \eta)$.

Solution. En vertu du corollaire 3 on a

$$(F, H) = [I\,dH, I\,dF] = [\text{grad}\, H, \text{grad}\, F] = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\frac{\partial F}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial F}{\partial p_i}$$

(nous nous servons du fait que I est symplectique et que dans la base (p, q) sa matrice est $\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$.

Exercice 2. Calculer le crochet de Poisson des fonctions coordonnées p_i, q_i.

Solution. Les gradients des fonctions coordonnées engendrent une « base symplectique »: leurs produits scalaires gauches valent

$$(p_i, p_j) = (p_i, q_j) = (q_i, q_j) = 0, \quad (q_i, p_i) = -(p_i, q_i) = 1.$$

Exercice 3. Montrer que *l'application* $A : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$, $(p, q) \to (P(p, q), Q(p, q))$ *est canonique si et seulement si les crochets de Poisson de deux fonctions quelconques en les variables (p, q) et (P, Q) se confondent:*

$$(F, H)_{p,q} = \frac{\partial H}{\partial p}\frac{\partial F}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial F}{\partial p} = \frac{\partial H}{\partial P}\frac{\partial F}{\partial Q} - \frac{\partial H}{\partial Q}\frac{\partial F}{\partial P} = (F, H)_{P,Q}.$$

Solution. Supposons que A est canonique. Alors les structures symplectiques $dp \wedge dq$ et $dP \wedge dQ$ se confondent. Or la définition du crochet de Poisson est intrinsèquement liée à la structure symplectique et non aux coordonnées. Donc

$$(F, H)_{p,q} = (F, H) = (F, H)_{P,Q}.$$

Réciproque. Supposons que le crochet de Poisson $(P_i, Q_i)_{p,q}$ est de la forme standard de l'exercice 2. Il est alors évident que $dP \wedge dQ = dp \wedge dq$, i.e. l'application A est canonique.

Exercice 4. Montrer que le *crochet de Poisson d'un produit se calcule d'après la règle de Leibniz*:

$$(F_1 F_2, H) = F_1 (F_2, H) + F_2 (F_1, H).$$

I n d i c a t i o n. Le crochet de Poisson (F_1F_2, H) est la dérivée du produit F_1F_2 suivant la direction du champ $I\,dH$.

B. Identité de Jacobi.

Théorème. *Les crochets de Poisson de trois fonctions A, B, C vérifient l'identité de Jacobi:*

$$((A, B), C) + ((B, C), A) + ((C, A), B) = 0.$$

Corollaire. Théorème de Poisson. *Le crochet de Poisson de deux intégrales premières (F_1, F_2) d'un système de hamiltonien H est de nouveau une intégrale première.*

D é m o n s t r a t i o n d u c o r o l l a i r e. L'identité de Jacobi entraîne

$$((F_1, F_2), H) = (F_1, (F_2, H)) + (F_2, (H, F_1)) = 0 + 0,$$

c.q.f.d.

Si donc l'on connaît deux intégrales premières, on peut par simple calcul déduire une troisième, une quatrième, etc. En fait, certaines des intégrales premières trouvées peuvent se confondre puisque sur M^{2n} il ne peut pas exister plus de $2n$ fonctions indépendantes. Parfois on peut trouver une fonction qui dépende des anciennes ou bien une constante, par exemple zéro. Mais on arrive quelquefois à trouver une intégrale nouvelle.

E x e r c i c e. Calculer les crochets de Poisson des composantes p_1, p_2, p_3, M_1, M_2, M_3 des vecteurs impulsion et moment cinétique d'un système mécanique.

Réponse. $(M_1, M_2) = M_3$, $(M_1, p_1) = 0$, $(M_1, p_2) = p_3$, $(M_1, p_3) = -p_2$. D'où résulte le

Théorème. *Si dans un système mécanique deux composantes, M_1 et M_2, du moment cinétique sont invariantes, alors la troisième composante, M_3, l'est également.*

D é m o n s t r a t i o n d e l ' i d e n t i t é d e J a c o b i. Considérons la somme

$$((A, B), C) + ((B, C), A) + ((C, A), B).$$

Cette somme est une « combinaison linéaire des dérivées partielles secondes » de fonctions. Calculons les membres qui contiennent les dérivées secondes de A :

$$((A, B), C) + ((C, A), B) = (L_C L_B - L_B L_C)\,A,$$

où L_ξ est la différentiation suivant la direction de ξ et B, C des champs hamiltoniens de hamiltoniens B, C.

Or en vertu du lemme 2, § 39, le commutateur de différentiation $L_C L_B - L_B L_C$ est un opérateur différentiel du *premier* ordre. Donc aucune dérivée seconde de A ne peut figurer dans notre somme. Il en est de même des dérivées secondes de B et C. Donc la somme est nulle, c.q.f.d.

Corollaire 5. *Soient* $\boldsymbol{B}$, $\boldsymbol{C}$ *des champs hamiltoniens de hamiltoniens* B *et* C. *Considérons le crochet de Poisson des champs de vecteurs* $\boldsymbol{B}$ *et* $\boldsymbol{C}$. *C'est un champ de vecteurs hamiltonien et sa fonction de Hamilton est égale au crochet de Poison* (B, C).

Démonstration. Posons $(B, C) = D$. L'identité de Jacobi peut se mettre sous la forme

$$(A, D) = ((A, B), C) - ((A, C), B),$$

$$L_D = L_C L_B - L_B L_C, \; L_D = L_{[B, C]},$$

c.q.f.d.

C. Algèbres de Lie de champs hamiltoniens, de hamiltoniens et d'intégrales premières. Un sous-espace vectoriel d'une algèbre de Lie est une *sous-algèbre* s'il contient le commutateur de deux quelconques de ses éléments. Une sous-algèbre de Lie est elle-même une algèbre de Lie. Le corollaire précédent englobe notamment le

Corollaire 6. *Les champs de vecteurs hamiltoniens sur une variété symplectique engendrent une sous-algèbre de l'algèbre de Lie de tous les champs.*

Le théorème de Poisson relatif aux intégrales premières peut encore s'énoncer:

Corollaire 7. *Les intégrales premières du flot hamiltonien engendrent une sous-algèbre de l'algèbre de Lie de toutes les fonctions.*

L'algèbre de Lie des hamiltoniens peut être naturellement appliquée sur l'algèbre de Lie des champs de vecteurs hamiltoniens. Pour cela à chaque fonction H on associe le champ de vecteurs hamiltonien $\boldsymbol{H}$ de fonction H.

Corollaire 8. *L'application de l'algèbre de Lie des fonctions sur l'algèbre de Lie des champs hamiltoniens est un homomorphisme d'algèbres dont le noyau est composé de fonctions localement constantes. Si* M *est connexe, alors ce noyau est de dimension* 1 *et se compose de constantes.*

Notre application est linéaire. Le corollaire 5 affirme qu'elle envoie un crochet de Poisson de fonctions dans un crochet de Poisson de champs de vecteurs. Le noyau est composé de fonctions H telles que $I\, dH \equiv 0$. Comme I est un isomorphisme, $dH \equiv 0$, $H = \text{const}$, c.q.f.d.

Corollaire 9. *Une condition nécessaire et suffisante pour que les flots de hamiltoniens* H_1 *et* H_2 *commutent est que le crochet de Poisson des fonctions* H_1 *et* H_2 *soit (localement) constant.*

D'après le théorème du point E du § 39, une condition nécessaire et suffisante est que $[\boldsymbol{H}_1, \boldsymbol{H}_2] \equiv 0$, or en vertu du corollaire 8 cette condition équivaut à $d\,(H_1, H_2) \equiv 0$.

Nous avons donc obtenu une nouvelle généralisation du théorème de E. Noether: *si l'on connaît un flot commutant avec le flot étudié, on peut construire une intégrale première.*

D. Champs de vecteurs localement hamiltoniens. Soit (M, ω^2) une variété symplectique, $g^t: M \to M$ un groupe à un paramètre de difféomorphismes conservant la structure symplectique. g^t sera-t-il flot hamiltonien?

Exemple. Soit M un tore T^2 de dimension 2 dont tout point est défini par un couple de coordonnées (p, q) modulo 1. Soit ω^2 l'élément usuel d'aire $dp \wedge dq$. Considérons la famille de translations $g^t(p, q) = (p + t, q)$ (fig. 173).

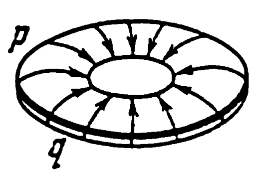


Fig. 173. Champ localement hamiltonien sur un tore.

Les applications g^t conservent la structure symplectique (i.e. l'aire). Peut-on associer le champ de vecteurs correspondant ($\dot{p} = 1$, $\dot{q} = 0$) à une fonction de Hamilton? Si $\dot{p} = -\partial H/\partial q$, $\dot{q} = \partial H/\partial p$, nous aurions eu $\partial H/\partial p = 0$, $\partial H/\partial q = -1$, i.e. $H = -q + C$. Or, q n'est qu'une coordonnée *locale* sur T^2; il n'existe pas d'application $H: T^2 \to \mathbf{R}$ telle que $\partial H/\partial p = 0$, $\partial H/\partial q = -1$.

Donc g^t n'est pas un flot hamiltonien.

Définition. On appelle *champ de vecteurs localement hamiltonien* sur une variété symplectique (M, ω^2) un champ de vecteurs $I\omega^1$, où ω^1 est une 1-forme fermée dans M.

La 1-forme localement fermée ω^1 est différentielle d'une fonction: $\omega^1 = dH$. Cependant si l'on essaie de prolonger la fonction H sur la variété M tout entière, on risque d'obtenir un « hamiltonien multivoque ». Car une 1-forme fermée dans une variété multiplement connexe n'est pas forcément différentielle (exemple: la forme dq sur T^2).

Le flot défini localement par un champ de vecteurs hamiltonien s'appelle *flot localement hamiltonien.*

Exercice. Montrer que *pour qu'un groupe à un paramètre de difféomorphismes d'une variété symplectique conserve la structure symplectique, il faut et il suffit qu'il soit flot localement hamiltonien.*

Indication. Cf. § 38, A.

Exercice. Montrer que *dans l'espace symplectique* $\mathbf{R}^{2n}$ *tout groupe à un paramètre de difféomorphismes canoniques (conservant* $dp \wedge dq$*) est toujours flot hamiltonien.*

Indication. Toute 1-forme fermée dans $\mathbf{R}^{2n}$ est différentielle d'une fonction.

Exercice. Montrer que les *champs de vecteurs localement hamiltoniens engendrent une sous-algèbre de l'algèbre de Lie de tous les champs et que le crochet de Poisson de deux champs localement hamiltoniens est un champ hamiltonien dont le hamiltonien est univoquement* *) *défini par la formule* $H = \omega^2(\xi, \eta)$, *où* ξ *et* η *sont les vecteurs des champs donnés.*

Donc les champs hamiltoniens engendrent un idéal dans l'algèbre de Lie des champs localement hamiltoniens.

---

*) Et non pas à une constante près.

§ 41. Géométrie symplectique

Une structure euclidienne d'un espace vectoriel est définie par une forme bilinéaire symétrique et une structure symplectique par une forme antisymétrique. La géométrie d'un espace symplectique ne ressemble pas à la géométrie euclidienne même si elle s'en rapproche par de nombreux traits.

A. Espace vectoriel symplectique. Soit $\mathbf{R}^{2n}$ un espace vectoriel de dimension paire.

D é f i n i t i o n. *Une structure symplectique linéaire sur* $\mathbf{R}^{2n}$ est une 2-forme régulière *), bilinéaire et antisymétrique sur $\mathbf{R}^{2n}$. Cette forme est appelée *produit scalaire gauche* et notée $[\xi, \eta] = - [\eta, \xi]$.

On appelle *espace vectoriel symplectique* un espace vectoriel réel $\mathbf{R}^{2n}$ sur lequel on a défini une structure symplectique.

E x e m p l e. Soit $(p_1, \ldots, p_n ; q_1, \ldots, q_n)$ des fonctions coordonnées sur $\mathbf{R}^{2n}$ et ω^2 une forme

$$\omega^2 = p_1 \wedge q_1 + \ldots + p_n \wedge q_n.$$

Comme cette forme est régulière et antisymétrique, on peut la prendre pour produit scalaire gauche: $[\xi, \eta] = \omega^2 (\xi, \eta)$. Donc l'espace arithmétique $\mathbf{R}^{2n} = \{(p, q)\}$ est muni d'une structure symplectique dite *standard*. Dans la structure symplectique standard le produit scalaire gauche de deux vecteurs ξ, η est égal à la somme des projections des aires orientées du parallélogramme (ξ, η) sur les n plans arithmétiques (p_i, q_i).

On dit que deux vecteurs ξ, η sont orthogonaux gauches et l'on note $\xi \angle \eta$ si leur produit scalaire gauche est nul.

E x e r c i c e. Montrer que $\xi \angle \xi$: *chaque vecteur est self-orthogonal gauche.*

L'ensemble de tous les vecteurs orthogonaux gauches à un vecteur donné η est appelé *orthocomplément gauche* de η.

E x e r c i c e. Montrer que l'orthocomplément gauche de η est un hyperplan de dimension $2n - 1$ contenant η.

I n d i c a t i o n. Si tous les vecteurs étaient orthogonaux gauches à η, la forme [,] serait dégénérée.

B. Base symplectique. Dans une base convenable (qui doit être orthonormée) une structure euclidienne est définie par un produit scalaire d'un type spécial. De la même façon une structure symplectique prend la forme standard indiquée plus haut dans une base adéquate.

E x e r c i c e. Trouver les produits scalaires gauches des vecteurs unitaires e_{p_i}, e_{q_i} $(i = 1, \ldots, n)$ de l'exemple cité plus haut.

---

*)Une 2-forme [,] sur $\mathbf{R}^{2n}$ est régulière si $([\xi, \eta] = 0 \ \forall \eta) \Rightarrow (\xi = 0)$.

S o l u t i o n. La définition de $p_1 \wedge q_1 + \ldots + p_n \wedge q_n$ entraîne les relations

$$[e_{p_i}, e_{p_j}] = [e_{p_i}, e_{q_j}] = [e_{q_i}, e_{q_j}] = 0, \quad [e_{p_i}, e_{q_i}] = 1. \tag{1}$$

Revenons maintenant à l'espace symplectique $\mathbf{R}^{2n}$.

D é f i n i t i o n. On appelle *base symplectique* $2n$ vecteurs e_{p_i}, e_{q_i} $(i = 1, \ldots, n)$ dont les produits scalaires gauches sont de la forme (1).

En d'autres termes, *tout vecteur unitaire est orthogonal gauche aux autres, sauf à un qui est son conjugué ; le produit de vecteurs conjugués est égal à* ± 1.

Théorème. *Dans tout espace symplectique il existe une base symplectique. Bien plus tout vecteur* e *non nul peut être pris pour premier vecteur de base.*

Ce théorème se démontre pratiquement de la même façon que son analogue euclidien.

Comme le vecteur e n'est pas nul, il existe un vecteur f qui ne lui est pas orthogonal gauche (la forme [,] est régulière). On peut choisir le module de f tel que le produit scalaire gauche de e et f soit égal à l'unité. Le théorème est démontré pour $n = 1$.

Si $n > 1$, considérons l'orthocomplément gauche D (fig. 174) du couple de vecteurs e, f. D est l'intersection des orthocompléments

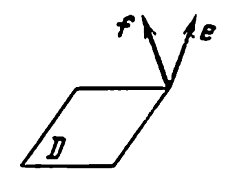



Fig. 174. Complémentaire orthogonal gauche.

gauches de e et de f. Ces deux sous-espaces de dimension $2n - 1$ ne sont pas confondus puisque e n'est pas contenu dans l'orthocomplément gauche de f. Donc leur intersection D a une dimension paire $2n - 2$.

Montrons que D est un sous-espace symplectique de $\mathbf{R}^{2n}$, i.e. le produit scalaire gauche [,] sur D n'est pas dégénéré. En effet, si un vecteur $\xi \in D$ était orthogonal gauche à l'espace D tout entier, alors il serait antiorthogonal à $\mathbf{R}^{2n}$ (puisqu'il est orthogonal gauche à e et f), ce qui contredit la régularité de la forme [,] sur $\mathbf{R}^{2n}$. Donc D^{2n-2} est symplectique.

Si maintenant à une base symplectique de D^{2n-2} on ajoute les vecteurs e et f, on obtient une base symplectique de $\mathbf{R}^{2n}$, et la démonstration du théorème s'achève par récurrence sur n.

Corollaire. *Tous les espaces symplectiques de même dimension sont isomorphes.*

Si l'on prend les vecteurs d'une base symplectique pour vecteurs unitaires, on obtient un système de coordonnées p_i, q_i dans lequel la forme [,] s'écrit $p_1 \wedge q_1 + \ldots + p_n \wedge q_n$. Un tel système de coordonnées est dit *symplectique.*

C. Groupe symplectique. A une structure euclidienne est lié un groupe orthogonal d'applications linéaires, conservant cette structure. Dans un espace symplectique ce rôle est assumé par le groupe symplectique.

Définition. Une transformation linéaire $S: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ d'un espace symplectique $\mathbf{R}^{2n}$ sur lui-même est *symplectique* si elle conserve le produit scalaire gauche

$$[S\xi, S\eta] = [\xi, \eta] \ \forall \xi, \eta \in \mathbf{R}^{2n}.$$

L'ensemble de toutes les transformations symplectiques de $\mathbf{R}^{2n}$ s'appelle *groupe symplectique* et se note Sp $(2n)$.

Que le produit de deux transformations symplectiques soit symplectique est chose évidente. Pour justifier le terme « groupe symplectique » il faut démontrer seulement qu'une transformation symplectique est régulière, et il sera alors clair que son inverse est également symplectique.

Exercice. Montrer que le groupe Sp (2) est isomorphe au groupe des matrices réelles d'ordre deux et de déterminant 1 et homéomorphe à une chambre à air.

Théorème. *Une transformation $S: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ d'un espace symplectique standard (p, q) est symplectique si et seulement si elle est linéaire et canonique, i.e. conserve la 2-forme différentielle*

$$\omega^2 = dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n.$$

Démonstration. Si l'on identifie naturellement à $\mathbf{R}^{2n}$ un espace qui lui est tangent, la 2-forme ω^2 se transforme en [,].

Corollaire. *Le déterminant de toute transformation symplectique est égal à l'unité.*

Démonstration. Nous savons déjà (§ 38, B) que les transformations canoniques préservent les puissances extérieures de la forme ω^2. Or sa n-ème puissance extérieure est (à un facteur constant près) l'élément de volume de $\mathbf{R}^{2n}$. Donc les transformations symplectiques S de l'espace standard $\mathbf{R}^{2n} = \{(p, q)\}$ conservent l'élément de volume et det $S = 1$.

Cependant comme toute structure symplectique linéaire prend sa forme standard dans un système de coordonnées symplectique, le déterminant de toute transformation symplectique d'un espace symplectique quelconque est égal à l'unité, c.q.f.d.

Théorème. *Une transformation linéaire S: $\mathbf{R}^{2n} \to \mathbf{R}^{2n}$ est symplectique si et seulement si elle envoie une certaine base symplectique (et par conséquent une base quelconque) dans une base symplectique.*

D é m o n s t r a t i o n. Le produit scalaire gauche de deux combinaisons linéaires quelconques de vecteurs unitaires s'exprime en fonction des produits scalaires gauches des vecteurs unitaires. Si la transformation ne change pas les produits scalaires gauches des vecteurs unitaires, *a fortiori* elle ne changera pas les produits scalaires gauches de vecteurs quelconques, c.q.f.d.

D. Plan dans un espace symplectique. Dans un espace euclidien tous les plans sont équivalents, i.e. chacun d'entre eux peut être transporté dans un autre par un mouvement.

Voyons s'il en est de même dans un espace vectoriel symplectique.

E x e r c i c e. Montrer qu'un vecteur non nul d'un espace symplectique peut être envoyé dans n'importe quel autre vecteur non nul par une transformation symplectique.

E x e r c i c e. Montrer que par une transformation symplectique d'un 2-plan donné *on ne peut obtenir tous les 2-plans* d'un espace symplectique $\mathbf{R}^{2n}$, $n > 1$.

I n d i c a t i o n. Etudier les plans (p_1, p_2) et (p_1, q_1).

D é f i n i t i o n. Un k-plan d'un espace symplectique est *isotrope* *) (ou nul) s'il est self-orthogonal gauche, i.e. le produit scalaire gauche de deux vecteurs quelconques de ce plan est nul.

E x e m p l e. Le plan arithmétique $(p_1, \ldots, p_k)$ est isotrope dans le système de coordonnées symplectique $\boldsymbol{p}$, $\boldsymbol{q}$. (Prouvez-le!)

E x e r c i c e. Montrer que tout 2-plan non isotrope peut être envoyé dans un 2-plan quelconque non isotrope par une transformation symplectique.

Lorsqu'on effectue des calculs en géométrie symplectique, il est souvent utile de munir l'espace symplectique encore d'une structure euclidienne. Fixons un système de coordonnées symplectique $\boldsymbol{p}$, $\boldsymbol{q}$ et introduisons une structure euclidienne moyennant le produit scalaire de coordonnées

$$(\boldsymbol{x}, \boldsymbol{x}) = \sum p_i^2 + q_i^2, \quad \text{où} \quad \boldsymbol{x} = \sum p_i e_{p_i} + q_i e_{q_i}.$$

La base symplectique $\boldsymbol{e_p}$, $\boldsymbol{e_q}$ est orthonormée dans cette structure euclidienne. Le produit scalaire gauche comme toute forme bilinéaire s'exprime en fonction du produit scalaire:

$$[\xi, \eta] = (I\xi, \eta), \tag{2}$$

où $I: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ est un opérateur. L'antisymétrie du produit scalaire gauche entraîne celle de l'opérateur I.

*) Si $k = n$, on les appelle *lagrangiens*.

Exercice. Trouver la matrice de l'opérateur I dans la base symplectique e_{p_i}, e_{q_i}.

Réponse. $\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$, où E est la matrice unité d'ordre n.

Donc pour $n = 1$ (sur le plan p, q) I est tout simplement une rotation de 90°, dans le cas général, I est une rotation de 90° dans chacun des n plans p_i, q_i.

Exercice. *Montrer que l'opérateur I est symplectique et que $I^2 = -E_{2n}$.*

La structure euclidienne et l'opérateur I sont d'un usage commode même s'ils sont intrinsèquement liés à l'espace symplectique.

La formule (2) entraîne aussitôt le

Théorème. *Un plan π d'un espace symplectique est isotrope si et seulement si le plan $I\pi$ est orthogonal à π.*

On remarquera que les plans π et $I\pi$ ont même dimension puisque I est régulier. D'où le

Corollaire. *La dimension d'un plan isotrope de $\mathbf{R}^{2n}$ n'est pas supérieure à n.*

En effet deux plans π et $I\pi$ de dimension k dans $\mathbf{R}^{2n}$ ne peuvent être orthogonaux si $k > n$.

Voyons un peu plus en détail les n-plans isotropes de l'espace symplectique arithmétique $\mathbf{R}^{2n}$. Un tel plan est le p-plan arithmétique (i.e. le plan $q = 0$). Dans $\mathbf{R}^{2n} = \{(p, q)\}$ il existe C_{2n}^n plans arithmétiques de dimension n.

Exercice. Montrer que parmi les C_{2n}^n plans arithmétiques de dimension n il en existe exactement 2^n isotropes. Plus précisément à chacune des 2^n partitions de l'ensemble $(1, \ldots, n)$ en deux parties $(i_1, \ldots, i_k)$ et. $(j_1, \ldots, j_{n-k})$ correspond un plan arithmétique isotrope $p_{i_1}, \ldots, p_{i_k}, q_{j_1}, \ldots, q_{j_{n-k}}$.

Pour étudier les fonctions génératrices des transformations canoniques nous aurons besoin du

Théorème. *Tout n-plan isotrope π de l'espace symplectique arithmétique $\mathbf{R}^{2n}$ est transversal *) au moins à l'un des 2^n plans arithmétiques isotropes.*

Démonstration. Soit P un plan isotrope $p_1, \ldots, p_n$ (fig. 175). Considérons l'intersection

$$\tau = \pi \cap P.$$

Supposons que τ est de dimension k, $0 \leqslant k \leqslant n$. Comme tout sous-espace de dimension k de l'espace P de dimension n, le plan τ est

*) Deux sous-espaces L_1 et L_2 d'un espace vectoriel L sont transversaux si $L_1 + L_2 = L$. Deux plans de dimension n de $\mathbf{R}^{2n}$ sont transversaux si et seulement si ils ne se coupent qu'au point O.

transversal au moins à l'un des $(n-k)$-plans de P, qui pour fixer les idées sera le plan

$$\eta = (p_{i_1}, \ldots, p_{i_{n-k}});\ \tau + \eta = P,\ \tau \cap \eta = 0.$$

Construisons un n-plan arithmétique isotrope

$$\sigma = (p_{i_1}, \ldots, p_{i_{n-k}}, q_{j_1}, \ldots, q_{j_k}),\ \eta = \sigma \cap P,$$

et montrons que π est transversal à σ:

$$\pi \cap \sigma = 0.$$

On a en effet

$$\left.\begin{array}{l} \tau \subset \pi,\ \pi \perp \pi \Rightarrow \tau \perp \pi \\ \eta \subset \sigma,\ \sigma \perp \sigma \Rightarrow \eta \perp \sigma \end{array}\right\} \Rightarrow (\tau + \eta) \perp (\pi \cap \sigma) \Rightarrow P \perp (\pi \cap \sigma).$$

Or P est un n-plan isotrope. Donc tout vecteur orthogonal gauche à P

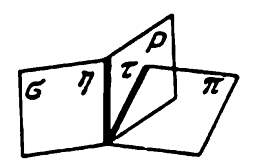


Fig. 175. Construction d'un plan arithmétique σ transversal à un plan donné π.

appartient à P (voir le corollaire plus haut). Et $(\pi \cap \sigma) \subset P$. Finalement

$$\pi \cap \sigma = (\pi \cap P) \cap (\sigma \cap P) = \tau \cap \eta = 0,$$

c.q.f.d.

Exercice. Soient π_1, π_2 deux k-plans d'un espace symplectique $\mathbf{R}^{2n}$. Peut-on toujours envoyer π_1 dans π_2 par une transformation symplectique? Combien existe-t-il de classes de plans non transformables l'un dans l'autre?

Réponse. $[k/2]+1$ si $k \leqslant n$; $\left[\frac{2n-k}{2}\right]+1$ si $k \geqslant n$.

E. Structure symplectique et structure complexe. Comme $I^2 = -E$, outre la structure symplectique $[,]$ et la structure euclidienne $(,)$ on peut munir notre espace $\mathbf{R}^{2n}$ d'une structure complexe en définissant le produit par $i = \sqrt{-1}$ comme l'effet de I. L'espace $\mathbf{R}^{2n}$ s'identifie alors à un espace complexe $\mathbf{C}^n$ ou si l'on veut à un espace arithmétique muni des coordonnées $z_k = p_k + iq_k$.

Les transformations linéaires de $\mathbf{R}^{2n}$ qui conservent la structure euclidienne forment le *groupe orthogonal* $O(n)$; celles qui conservent la structure complexe, le groupe linéaire complexe $GL(n, \mathbf{C})$.

Exercice. Montrer que *les transformations à la fois orthogonales et symplectiques, complexes et orthogonales, symplectiques et complexes sont respectivement complexes, symplectiques et orthogonales*, de sorte que les intersections de

deux des trois groupes sont égales à l'intersection des trois groupes:

$O(2n) \cap Sp(2n) = Sp(2n) \cap GL(n, C) = GL(n, C) \cap O(2n).$

Cette intersection s'appelle *groupe unitaire* $U(n)$.

Les transformations unitaires conservent le produit scalaire hermitien $(\xi, \eta) + i[\xi, \eta]$; les produits scalaire et scalaire gauche dans R^{2n} en constituent la partie réelle et la partie imaginaire.

§ 42. Résonance paramétrique dans les systèmes à plusieurs degrés de liberté

En étudiant les systèmes oscillatoires à paramètres périodiques (cf. § 25) nous avons vu que la résonance paramétrique dépendait du comportement des valeurs propres d'une certaine transformation linéaire (une « application en une période »). Cette dépendance consistait en ce que les états d'équilibre du système à paramètres périodiques étaient stables si les valeurs propres de l'application en une période étaient inférieures en module à l'unité, et instables si l'une d'elles au moins était supérieure à l'unité.

L'application en une période déduite d'un système d'équations de Hamilton à coefficients périodiques est symplectique. L'étude de la résonance paramétrique dans les systèmes à un degré de liberté, faite au § 25, se fondait sur l'analyse du comportement des valeurs propres de transformations symplectiques du plan.

Dans le présent paragraphe l'on se propose de procéder à une analyse analogue du comportement des valeurs propres des transformations symplectiques linéaires d'un espace des phases de dimension quelconque. Cette analyse (qui est due à M. Krein) s'applique à l'étude des conditions d'existence de la résonance paramétrique dans les systèmes mécaniques à plusieurs degrés de liberté.

A. Matrices symplectiques. Considérons une transformation linéaire d'un espace symplectique $S: R^{2n} \to R^{2n}$. Soit $p_1, \ldots, p_n$; $q_1, \ldots, q_n$ un système de coordonnées symplectique. Dans ce système la transformation envisagée est définie par une matrice S.

Théorème. *Pour qu'une transformation soit symplectique il est nécessaire et suffisant que dans le système de coordonnées* (p, q) *sa matrice vérifie la relation*

$$S'IS = I,$$

où $I = \begin{pmatrix} 0 & -E' \\ E & 0 \end{pmatrix}$ *et* S' *est la transposée de* S.

D é m o n s t r a t i o n. Si l'on fait intervenir l'opérateur I et le produit scalaire, la condition de symplecticité ($[S\xi, S\eta] = [\xi, \eta]$ pour tous les ξ et η s'écrit

$$(IS\xi, S\eta) = (I\xi, \eta) \ \forall \ \xi, \eta,$$

ou encore

$$(S'IS\xi, \eta) = (I\xi, \eta) \ \forall \ \xi, \eta,$$

c.q.f.d.

B. Symétrie du spectre d'une transformation symplectique.

Théorème. *Le polynôme caractéristique d'une transformation symplectique*

$$p(\lambda) = \det|S - \lambda E|$$

est récurrent *), *i.e.* $p(\lambda) = \lambda^{2n} p(1/\lambda)$.

Démonstration. Nous nous servirons du fait que $\det S = \det I = 1$, $I^2 = -E$ et $\det A' = \det A$. D'après le théorème précédent $S = -IS'^{-1}I$. Donc

$$p(\lambda) = \det(S - \lambda E) = \det(-IS'^{-1}I - \lambda E) =$$
$$= \det(-S'^{-1} + \lambda E) = \det(-E + \lambda S) = \lambda^{2n} \det\left(S - \frac{1}{\lambda}E\right) =$$
$$= \lambda^{2n} p\left(\frac{1}{\lambda}\right), \text{ c.q.f.d.}$$

Corollaire. *Si* λ *est valeur propre d'une transformation symplectique, alors* $1/\lambda$ *l'est également.*

D'autre part, le polynôme caractéristique est réel; donc si λ est une valeur propre complexe, alors $\bar{\lambda}$ est également une valeur propre complexe distincte de λ.

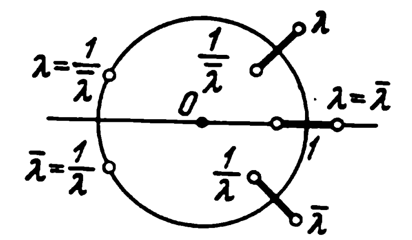


Fig. 176. Répartition des valeurs propres d'une application symplectique.

Il s'ensuit que toutes les racines λ du polynôme caractéristique sont symétriques par rapport à l'axe réel et au cercle unité (fig. 176). Elles se répartissent en quaternes

$$\lambda, \bar{\lambda}, \frac{1}{\lambda}, \frac{1}{\bar{\lambda}} \quad (|\lambda| \neq 1, \operatorname{Im}\lambda \neq 0)$$

et en couples situés sur l'axe réel

$$\lambda = \bar{\lambda}, \ \frac{1}{\lambda} = \frac{1}{\bar{\lambda}}$$

ou sur le cercle unité:

$$\lambda = \frac{1}{\bar{\lambda}}, \ \bar{\lambda} = \frac{1}{\lambda}.$$

Il est aisé d'établir que les quatre points d'un quaterne (ou les deux points d'un couple) ont même multiplicité.

*) Un polynôme $a_0x^m + a_1x^{m-1} + \ldots + a_m$ est récurrent si ses coefficients sont symétriques, i. e. $a_0 = a_m$, $a_1 = a_{m-1}, \ldots$

C. Stabilité.

Définition. Une transformation S est *stable* si

$$\forall \varepsilon > 0 \ \exists \ \delta > 0 : |x| < \delta \Rightarrow |S^N x| < \varepsilon, \ \forall N > 0.$$

Exercice. Montrer qu'*une transformation symplectique S est instable si l'une au moins de ses valeurs propres n'est pas située sur le cercle unité.*

Indication. En vertu de la symétrie démontrée, si l'une au moins des valeurs propres de S n'est située sur le cercle unité, il en existe une autre $|\lambda| > 1$ qui est extérieure au cercle; dans le sous-espace invariant correspondant, S est le produit d'une homothétie par une rotation, i.e. une similitude.

Exercice. Montrer qu'*une transformation linéaire est stable si toutes ses valeurs propres sont distinctes et situées sur le cercle unité.*

Indication. Passer à une base propre.

Définition. Une transformation symplectique S est *fortement stable* si est stable toute transformation symplectique S_1 suffisamment proche *) de S.

Au § 25 nous avons établi que $S : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ est fortement stable si $\lambda_{1,2} = e^{\pm i\alpha}$, $\lambda_1 \neq \lambda_2$.

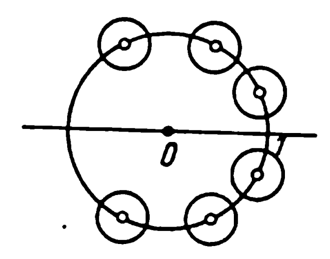


Fig. 177. Comportement des valeurs propres simples par une faible variation de la transformation symplectique.

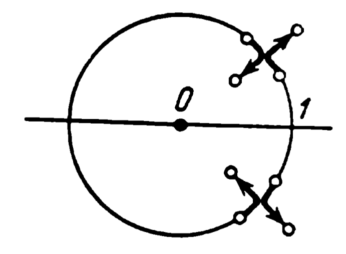


Fig. 178. Comportement des valeurs propres multiples par une faible variation de la transformation symplectique.

Théorème. *Si les $2n$ valeurs propres d'une transformation symplectique S sont distinctes et situées sur le cercle unité, alors la transformation S est fortement stable.*

Démonstration. Comprenons les $2n$ valeurs propres λ dans $2n$ voisinages disjoints et symétriques par rapport au cercle unité et à l'axe réel (fig. 177). Les $2n$ racines du polynôme caractéristique dépendent continûment des éléments de la matrice S. Donc, si une matrice S_1 est suffisamment proche de S, alors chacun des $2n$ voisinages des $2n$ points λ renferme une seule valeur propre λ_1 de la matrice S_1. Or si l'un des points, λ_1, n'était pas situé sur le cercle unité, mais, par exemple, à l'extérieur, d'après le corollaire

*) S_1 est « suffisamment proche » de S si les éléments de leurs matrices par rapport à une même base fixe présentent un écart inférieur à un nombre ε arbitrairement petit.

du théorème du pt. B, § 42, ce même voisinage contiendrait un autre point, λ_2, $|\lambda_2| < 1$, et le nombre total de racines serait supérieur à $2n$, ce qui est impossible.

Donc toutes les valeurs propres de S_1 sont situées sur le cercle unité et sont distinctes : S_1 est stable, c.q.f.d.

On peut dire qu'une valeur propre λ d'une transformation symplectique n'est susceptible de quitter le cercle unité qu'une fois entrée en collision avec une autre valeur propre (fig. 178) ; la collision a lieu entre des nombres conjugués complexes, et deux couples de racines situées sur le cercle donnent un quaterne (ou un couple de réels λ).

Des résultats du § 25 il vient que la résonance paramétrique se manifeste dans un système canonique linéaire à hamiltonien périodique lorsque précisément la transformation symplectique correspondante de l'espace des phases cesse d'être stable. D'après le théorème démontré ceci ne peut avoir lieu qu'en cas de collision de valeurs propres sur le cercle unité. En réalité, comme l'a souligné M. Krein, les collisions ne sont pas toutes dangereuses.

Il se trouve que les valeurs propres λ, $|\lambda| = 1$, se répartissent en deux classes : les *positives* et les *négatives*. Lorsque deux valeurs propres de même signe se rencontrent, elles « passent l'une à travers l'autre » et restent sur le cercle unité. Par contre, si elles sont de signes contraires, en général elles quittent toutes deux le cercle unité.

La théorie de M. Krein sort du cadre de notre ouvrage, aussi nous bornerons-nous à en énoncer les principaux résultats sous forme d'exercices.

Exercice. Soient λ, $\bar{\lambda}$ des valeurs propres simples (de multiplicité 1) d'une transformation symplectique S et $|\lambda| = 1$. Montrer que le 2-plan invariant π_λ correspondant à λ, $\bar{\lambda}$ n'est pas nul.

Indication. Soient ξ_1, ξ_2 des vecteurs propres complexes de S associés aux valeurs propres λ_1, λ_2. Si $\lambda_1\lambda_2 \neq 1$, les vecteurs ξ_1, ξ_2 sont orthogonaux gauches : $[\xi_1, \xi_2] = 0$.

Soit ξ un vecteur réel du plan π_λ, $\operatorname{Im} \lambda > 0$, $|\lambda| = 1$. La valeur propre λ est *positive* si $[S\xi, \xi] > 0$.

Exercice. Montrer que cette définition est correcte, i.e. ne dépend pas du choix du vecteur $\xi \neq 0$ sur le plan π_λ.

Indication. Si le plan π_λ contenait deux vecteurs orthogonaux gauches non parallèles, il serait nul.

De la même façon, une valeur propre λ, $|\lambda| = 1$, de multiplicité k est *définie positive* (resp. *négative*) si la forme quadratique $[S\xi, \xi]$ est définie positive (resp. négative) sur le sous-espace invariant de dimension $2k$ correspondant à λ, $\bar{\lambda}$.

Exercice. Montrer qu'une condition nécessaire et suffisante pour que S soit fortement stable est que chaque valeur propre λ soit située sur le cercle unité et soit définie positive ou négative.

Indication. La forme quadratique $[S\xi, \xi]$ est invariante par S.

§ 43. Atlas symplectique

Dans ce paragraphe on démontre le théorème de Darboux qui dit que toute variété symplectique possède des coordonnées locales p, q dans lesquelles la structure symplectique prend la forme excessivement simple: $\omega^2 = dp \wedge dq$.

A. Coordonnées symplectiques. Rappelons que dans la définition d'une variété participe la condition de compatibilité des cartes d'un atlas. Cette condition porte sur les applications $\varphi_i^{-1}\varphi_j$ de passage d'une carte à une autre. Les applications $\varphi_i^{-1}\varphi_j$ sont des applications de domaines de l'espace arithmétique.

D é f i n i t i o n. Un atlas d'une variété M^{2n} est dit *symplectique* si l'espace arithmétique $\mathbf{R}^{2n} = \{(p, q)\}$ est muni d'une structure symplectique standard $\omega^2 = dp \wedge dq$ et si l'on passe d'une carte à une autre par une transformation canonique *) (i.e. conservant ω^2), $\varphi_i^{-1}\varphi_j$.

E x e r c i c e. Montrer qu'un atlas symplectique définit une structure symplectique sur M^{2n}.

La réciproque est également vraie: toute variété symplectique possède un atlas symplectique. Ceci découle du théorème suivant.

B. Théorème de Darboux.

Théorème. *Soit ω^2 une 2-forme différentielle fermée régulière au voisinage d'un point x de l'espace $\mathbf{R}^{2n}$. Alors dans un voisinage du point x, on peut choisir un système de coordonnées locales $(p_1, \ldots, p_n;$ $q_1, \ldots, q_n)$ tel que la forme ω^2 prenne la forme standard*

$$\omega^2 = \sum_{i=1}^{n} dp_i \wedge dq_i.$$

Ce théorème permet de généraliser immédiatement à toutes les variétés symplectiques toute proposition de caractère local, invariante par les transformations canoniques et démontrée pour l'espace des phases standard $(\mathbf{R}^{2n}, \omega^2 = dp \wedge dq)$.

C. Construction des coordonnées p_1 et q_1. Pour première coordonnée, p_1, prenons une fonction linéaire non constante (on pourrait prendre une fonction différentiable quelconque dont la différentielle est non nulle au point x). Pour simplifier nous supposerons que $p_1(x) = 0$. Désignons par $P_1 = I\, dp_1$ le champ hamiltonien correspondant à la fonction p_1 (fig. 179). On remarquera que $P_1(x) \neq 0$. On peut donc faire passer par le point x un hyperplan N^{2n-1} ne contenant pas le vecteur $P_1(x)$ (au lieu de N^{2n-1} on aurait pu prendre une surface quelconque transversale à $P_1(x)$).

Considérons le flot hamiltonien P_1^t de hamiltonien p_1. Envisageons le temps t nécessaire pour aller de N au point $z = P_1^t y$ $(y \in N)$

*) On définit de façon analogue par exemple les variétés complexes analytiques: l'espace arithmétique doit être muni d'une structure complexe et le passage d'une carte à une autre doit être complexe et analytique.

par le flot P_1^t comme une fonction du point z. D'après les théorèmes de la théorie des équations différentielles ordinaires, cette fonction est définie et différentiable au voisinage du point $x \in \mathbf{R}^{2n}$. Désignons-la

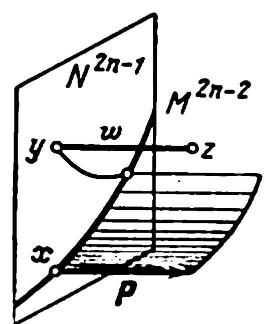


Fig. 179. Construction de coordonnées symplectiques.

par q_1. Remarquons que $q_1 = 0$ sur N et que la dérivée de la fonction q_1 suivant la direction du champ P_1 est égale à 1. Donc le *crochet de Poisson des fonctions q_1 et p_1 est égal* à 1 :

$$(q_1, p_1) \equiv 1.$$

D. Construction des coordonnées symplectiques par récurrence sur n. Si $n = 1$, la construction est achevée. Supposons que $n > 1$ et que le théorème de Darboux a été déjà démontré pour $\mathbf{R}^{2n-2}$.

Considérons l'ensemble M défini par les équations $q_1 = p_1 = 0$. Les différentielles dp_1 et dq_1 sont linéairement indépendantes au point x puisque $\omega^2 (I\,dp_1, I\,dq_1) = (q_1, p_1) \equiv 1$. Donc d'après le théorème des fonctions implicites, *au voisinage du point x l'ensemble M est une variété de dimension* $2n - 2$; désignons-la par M^{2n-2}.

Lemme. *La structure symplectique ω^2 sur $\mathbf{R}^{2n}$ définit une structure symplectique dans un voisinage du point x sur M^{2n-2}.*

D é m o n s t r a t i o n. Il nous faut simplement démontrer que ω^2 est régulière sur TM_x. Envisageons l'espace vectoriel symplectique $T\mathbf{R}^{2n}x$. Les vecteurs $P_1(x)$, $Q_1(x)$ des champs hamiltoniens de fonctions p_1 et q_1 appartiennent à $T\mathbf{R}^{2n}x$. Soit $\xi \in TM_x$. Les dérivées de p_1 et q_1 suivant la direction de ξ sont nulles. Donc $dp_1(\xi) = \omega^2(P_1, \xi) = 0$, $dq_1(\xi) = \omega^2(Q_1, \xi) = 0$ et TM_x est l'orthocomplément gauche de $P_1(x)$, $Q_1(x)$. En vertu du § 41, B, la forme ω^2 sur TM_x est régulière, c.q.f.d.

Par hypothèse de la récurrence, sur la variété symplectique $(M^{2n-2}, \omega^2|_M)$ existent des coordonnées symplectiques au voisinage du point x. Désignons-les par p_i, q_i $(i = 2, \ldots, n)$. Prolongeons les fonctions $p_2, \ldots, q_n$ dans le voisinage du point $x \in \mathbf{R}^{2n}$ de la manière suivante. Tout point z du voisinage de $x \in \mathbf{R}^{2n}$ peut être représenté de façon unique sous la forme $z = P_1^t Q_1^s w$, où $w \in M^{2n-2}$, et s et t sont des nombres petits. *Posons les valeurs des coordonnées $p_2, \ldots, q_n$ égales au point z et au point w* (fig. 179).

Les $2n$ fonctions $p_1, \ldots, p_n$; $q_1, \ldots, q_n$ construites forment un système de coordonnées local dans le voisinage de $x \in \mathbf{R}^{2n}$.

E. Démonstration de la symplecticité des coordonnées construites. Désignons par P_i^t, Q_i^s $(i = 1, \ldots, n)$ les flots hamiltoniens de fonctions p_i et q_i et par $\boldsymbol{P}_i$ et $\boldsymbol{Q}_i$ les champs de vecteurs correspondants. Calculons les crochets de Poisson des fonctions $p_1, \ldots, q_n$. Nous avons déjà vu au point C que $(q_1, p_1) \equiv 1$. Donc les flots P_1^t et Q_1^s commutent: $P_1^t Q_1^s = Q_1^s P_1^t$.

En se rappelant la définition des fonctions $p_2, \ldots, q_n$ on constate que chacune d'elles est invariante par les flots P_1^t et Q_1^s. Donc les *crochets de Poisson de p_1 et q_1 avec les $2n - 2$ fonctions p_i, q_i $(i > 1)$ sont nuls.*

L'application $P_1^t Q_1^s$ commute donc avec tous les $2n - 2$ flots $P_i^t Q_i^s$ $i > 1$. Par conséquent cette application laisse invariant chacun des $2n - 2$ champs de vecteurs $\boldsymbol{P}_i, \boldsymbol{Q}_i (i > 1)$. L'application $P_1^t Q_1^s$ conserve la structure symplectique ω^2, puisque les flots P_1^t et Q_1^s sont hamiltoniens. Donc la forme ω^2 prend aux points $z = P_1^t Q_1^s w \in \in \mathbf{R}^{2n}$ et $w \in M^{2n-2}$ les mêmes valeurs sur les vecteurs de deux quelconques des $2n - 2$ champs de vecteurs $\boldsymbol{P}_i$ et $\boldsymbol{Q}_i$ $(i > 1)$. Or ces valeurs sont égales à celles des crochets de Poisson des hamiltoniens correspondants. Donc les *valeurs du crochet de Poisson de deux quelconques des $2n-2$ coordonnées p_i, q_i $(i > 1)$ aux points z et w sont égales si $z = P_1^t Q_1^s w$.*

Les fonctions p_1 et q_1 sont intégrales premières de chacun des $2n-2$ flots P_i^t, Q_i^s $(i > 1)$. Donc chacun des $2n-2$ champs $\boldsymbol{P}_i$, $\boldsymbol{Q}_i$ $(i > 1)$ est tangent à la variété de niveau $p_1 = q_1 = 0$. Or cette variété n'est autre que M^{2n-2}. Donc chacun des $2n - 2$ champs $\boldsymbol{P}_i$, $\boldsymbol{Q}_i$ $(i > 1)$ est tangent à M^{2n-2}. Par conséquent ces champs sont hamiltoniens sur la variété symplectique $(M^{2n-2}, \omega^2|_M)$ et les hamiltoniens correspondants sont $p_i|_M$, $q_i|_M$ $(i > 1)$. Ainsi, *sur l'espace $(\mathbf{R}^{2n}, \omega^2)$ tout entier, le crochet de Poisson de deux quelconques des $2n - 2$ coordonnées p_i, q_i $(i > 1)$ considéré sur M^{2n-2} se confond avec le crochet de Poisson de ces coordonnées sur l'espace symplectique $(M^{2n-2}, \omega^2|_M)$.*

Or par hypothèse de la récurrence, les coordonnées sur M^{2n-2}, $(p_i|_M, q_i|_M ; i > 1)$ sont symplectiques. Donc sur l'espace $\mathbf{R}^{2n}$ tout entier le crochet de Poisson des coordonnées construites prend les valeurs standard

$$(p_i, p_j) \equiv (p_i, q_j) \equiv (q_i, q_j) \equiv 0, \ (q_i, p_i) \equiv 1.$$

Les crochets de Poisson des coordonnées $\boldsymbol{p}$, $\boldsymbol{q}$ sur $\mathbf{R}^{2n}$ sont de la même forme si $\omega^2 = \Sigma dp_i \wedge dq_i$. Or la forme bilinéaire ω^2 est définie par ses valeurs sur les couples de vecteurs unitaires. Donc les crochets de Poisson des fonctions coordonnées définissent ω^2 de façon unique. Et

$$\omega^2 = dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n.$$

Le théorème de Darboux est démontré.

CHAPITRE 9

FORMALISME CANONIQUE

Dans ce chapitre l'exposé se fera essentiellement dans le langage des coordonnées. L'appareil des fonctions génératrices des transformations canoniques développé par Hamilton et Jacobi est le plus puissant des procédés d'intégration des équations différentielles de la dynamique. Le flot hamiltonien est traité dans un espace de dimension impaire.

Ce chapitre est indépendant du précédent. Il renferme de nouvelles démonstrations de certains résultats du chapitre 8 et explique l'origine de la théorie des variétés symplectiques.

§ 44. Invariant intégral de Poincaré-Cartan

Dans ce paragraphe on étudie la géométrie d'une 1-forme sur un espace de dimension impaire.

A. Lemme d'hydrodynamique. Soit $\boldsymbol{v}$ un champ de vecteurs sur l'espace euclidien orienté $\mathbf{R}^3$, $\boldsymbol{r} = \operatorname{rot} \boldsymbol{v}$ le champ de son rotationnel. Les courbes intégrales de $\boldsymbol{r}$ sont appelées *caractéristiques*. Soit

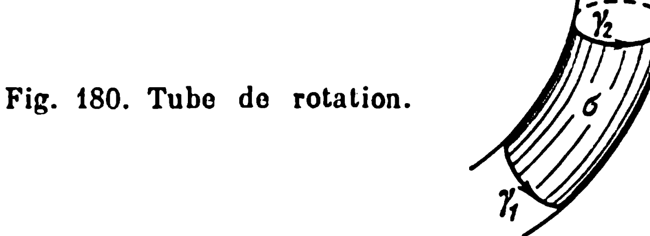


Fig. 180. Tube de rotation.

γ_1 une courbe fermée de $\mathbf{R}^3$ (fig. 180). Les caractéristiques passant par les points de γ_1 engendrent un *tube de caractéristiques* (ou *tube de rotationnel*).

Soit γ_2 une autre courbe englobant le même tube de caractéristiques de sorte que $\gamma_1 - \gamma_2 = \partial\sigma$, où σ est une 2-chaîne représentant une partie du tube de caractéristiques. On a le

Lemme de Stokes. *La circulation du champ v prend la même valeur le long des courbes γ_1 et γ_2:*

$$\oint_{\gamma_1} v\, dl = \oint_{\gamma_2} v\, dl.$$

D é m o n s t r a t i o n. D'après la formule de Stokes $\int_{\gamma_1} v dl - \int_{\gamma_2} v dl = \iint_{\sigma} \text{rot } v dn = 0$, puisque rot v est tangent au tube de caractéristiques, c.q.f.d.

B. Généralisation du lemme de Stokes. Le lemme de Stokes se généralise au cas d'une variété M^{2n+1} quelconque de dimension impaire (au lieu de $\mathbf{R}^3$). Pour formuler la généralisation du lemme passons des champs de vecteurs aux formes différentielles.

La circulation du champ v est l'intégrale de la 1-forme ω^1 ($\omega^1(\xi) = (v, \xi)$). Au rotationnel du champ v correspond la 2-forme $\omega^2 = d\omega^1$ ($d\omega^1(\xi, \eta) = (r, \xi, \eta)$). Cette formule montre qu'en chaque point il existe une direction (précisément la direction du rotationnel r (fig. 181)) telle que la circulation de v le long du bord de toute « aire infiniment petite » contenant r soit nulle:

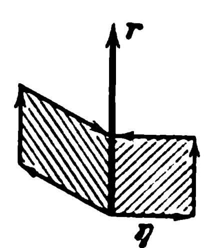


Fig. 181. Axe intrinsèquement lié à une 2-forme dans un espace de dimension impaire.

$$d\omega^1 (r, \eta) = 0 \ \forall \eta.$$

En effet $d\omega^1 (r, \eta) = (r, r, \eta) = 0$.

R e m a r q u e. Le passage de la 2-forme $\omega^2 = d\omega^1$ au champ du rotationnel r *n'est pas* une opération *invariante*: il dépend de la structure euclidienne de $\mathbf{R}^3$.

Cepfendant la *direction* du rotationnel r est intrinsèquement liée à la 2-rorme ω^2 (et partant à la 1-forme ω^1). En effet il est aisé de vérifier que si $r \neq 0$, alors la direction de r est définie de façon unique par la condition $\{\omega^2 (r, \eta) = 0 \ \forall \eta\}$.

La généralisation du lemme se base sur un fait algébrique qui dit que toute rotation d'un espace de dimension impaire possède un axe.

Lemme. *Soit ω^2 une 2-forme extérieure algébrique sur un espace vectoriel $\mathbf{R}^{2n+1}$ de dimension impaire. Il existe alors un vecteur $\xi \neq 0$ tel que*

$$\omega^2(\xi, \eta) = 0 \quad \forall \eta \in \mathbf{R}^{2n+1}.$$

D é m o n s t r a t i o n. La forme antisymétrique ω^2 est définie par une matrice antisymétrique A]

$$\omega^2(\xi, \eta) = (A\xi, \eta)$$

d'ordre impair $2n+1$. Le déterminant d'une telle matrice est nul puisque
$A' = -A$, $\det A = \det A' = \det(-A) = (-1)^{2n+1} \det A = -\det A$.
Donc le déterminant de A est nul. Cela veut dire que A possède un vecteur propre $\xi \neq 0$ associé à la valeur propre 0, c.q.f.d.

Un vecteur ξ tel que $\omega^2(\xi, \eta) \equiv 0$ pour tous les η est appelé *vecteur nul de la forme* ω^2. Tous les vecteurs nuls de ω^2 forment visiblement un sous-espace vectoriel. Une forme est dite *non singulière* si la dimension de cet espace est la plus petite possible (i.e. égale à 1 dans un espace $\mathbf{R}^{2n+1}$ de dimension impaire et à 0 dans un espace de dimension paire).

E x e r c i c e. Considérons sur un espace $\mathbf{R}^{2n}$ de dimension paire muni des coordonnées $p_1, \ldots, p_n; q_1, \ldots, q_n$ la 2-forme $\omega^2 = dp_1 \wedge dq_1 + \ldots + dp_n \wedge dq_n$. Montrer que la forme ω^2 est non singulière.

E x e r c i c e. Considérons sur un espace $\mathbf{R}^{2n+1}$ de dimension impaire muni des coordonnées $p_1, \ldots, p_n; q_1, \ldots, q_n; t$ la 2-forme $\omega^2 = \sum p_i\, dq_i - \omega^1\, dt$, où ω^1 est une 1-forme quelconque sur $\mathbf{R}^{2n+1}$. Montrer que la forme ω^2 est non singulière.

Si ω^2 est une forme *non singulière* sur $\mathbf{R}^{2n+1}$, tous les vecteurs nuls ξ de la forme ω^2 sont situés sur une même droite. Cette droite est intrinsèquement liée à la forme ω^2.

Soit maintenant M une variété différentiable de dimension impaire, ω^1 une 1-forme sur M. En vertu du lemme précédent, *en chaque point $x \in M$ il existe une direction (i.e. une droite $\{c\xi\}$ de l'espace tangent TM_x) telle que l'intégrale de ω^1 sur le bord d'une « aire infiniment petite contenant cette direction » soit nulle*:

$$d\omega^1(\xi, \eta) = 0 \quad \forall \eta \in TM_x.$$

Supposons par ailleurs que la 2-forme $d\omega^1$ est non singulière. Alors la direction de ξ est définie de façon unique. Nous appellerons cette direction « direction du rotationnel » de la forme ω^1.

Les courbes intégrales du champ de directions du rotationnel sont appelées *caractéristiques* de la forme ω^1.

Soit γ_1 une courbe fermée sur M. Les caractéristiques issues des points de γ_1 forment un tube. On a la

G é n é r a l i s a t i o n d u l e m m e d e S t o k e s: *L'intégrale de la 1-forme ω^1 prend la même valeur sur deux courbes γ_1 et γ_2 englobant un même tube de caractéristiques, i.e.* $\oint_{\gamma_1} \omega^1 = \oint_{\gamma_2} \omega^1$ *si $\gamma_1 - \gamma_2 = \partial\sigma$, où σ est un morceau de tube de caractéristiques.*

D é m o n s t r a t i o n. D'après la formule de Stokes

$$\oint_{\gamma_1} \omega^1 - \oint_{\gamma_2} \omega^1 = \oint_{\partial\sigma} \omega^1 = \int_{\sigma} d\omega^1.$$

Or $d\omega^1$ est nul sur un couple quelconque de vecteurs tangents au tube de caractéristiques. (En effet ces deux vecteurs sont situés dans un 2-plan passant par la direction du rotationnel, un plan sur lequel $d\omega^1$ est nul.)

Donc $\int_\sigma d\omega^1 = 0$, c.q.f.d.

C. Equations canoniques de Hamilton. Du lemme de Stokes découlent aussitôt les principaux théorèmes de la mécanique hamiltonienne.

Prenons pour M « l'espace des phases élargi $\mathbf{R}^{2n+1}$ » muni des

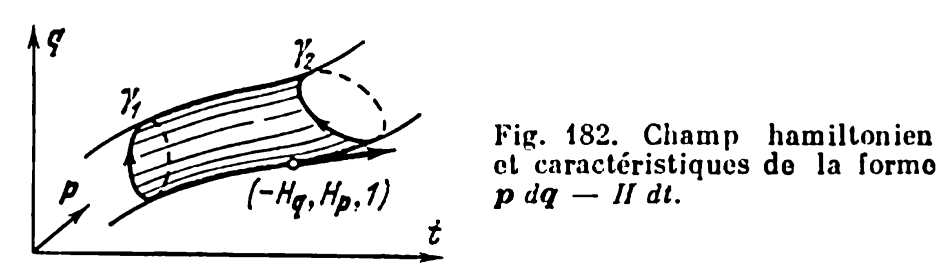


Fig. 182. Champ hamiltonien et caractéristiques de la forme $p\,dq - H\,dt$.

coordonnées $p_1, \ldots, p_n$; $q_1, \ldots, q_n$; t. Soit donnée une fonction $H = H(p, q, t)$. On peut alors composer *) la 1-forme

$$\omega^1 = p\,dq - H\,dt \quad (p\,dq = p_1 dq_1 + \ldots + p_n dq_n).$$

Appliquons à ω^1 le lemme de Stokes (fig. 182).

Théorème. *Les caractéristiques de la forme $\omega^1 = p\,dq - H\,dt$ sur l'espace des phases élargi $\{(p, q, t)\}$ de dimension $2n + 1$ se projettent de façon unique sur l'axe t, i.e. sont définies par des fonctions $p = p(t)$ et $q = q(t)$. Ces fonctions vérifient le système d'équations différentielles canoniques de hamiltonien H* :

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q}, \quad \frac{dq}{dt} = \frac{\partial H}{\partial p}. \tag{1}$$

En d'autres termes, les *caractéristiques de la forme $\omega^1 = p\,dq - H\,dt$ sont les trajectoires d'un flot dans l'espace des phases élargi, i.e. les courbes intégrales des équations canoniques* (1).

---

*) La forme ω^1 semble tomber des nues. Nous verrons dans la suite comment est venue l'idée d'envisager cette forme en optique.

D é m o n s t r a t i o n. La différentielle de la forme $p\,dq$ — — $H\,dt$ est

$$d\omega^1 = \sum_{i=1}^{n} \left(dp_i \wedge dq_i - \frac{\partial H}{\partial p_i} dp_i \wedge dt - \frac{\partial H}{\partial q_i} dq_i \wedge dt \right).$$

On voit donc que la matrice de la 2-forme $d\omega^1$ s'écrit dans les coordonnées p, q, t (vérifier!)

$$A = \begin{Vmatrix} 0 & -E & H_p \\ E & 0 & H_q \\ -H_p & -H_q & 0 \end{Vmatrix}, \quad \text{où} \quad E = \underbrace{\begin{Vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{Vmatrix}}_{n},$$

$$H_p = \frac{\partial H}{\partial p}, \quad H_q = \frac{\partial H}{\partial q}.$$

Le rang de cette matrice est égal à $2n$ (le mineur $2n \times 2n$ formé des éléments du coin supérieur gauche est non dégénéré), donc la 2-forme $d\omega^1$ est non singulière. Il est immédiat de vérifier que le vecteur $(-H_q, H_p, 1)$ est vecteur propre de la matrice A associé à la valeur propre 0 (vérifier!). Donc il définit la direction des caractéristiques de la forme $p\,dq - H dt$. Or le vecteur $(-H_q, H_p, 1)$ est justement le vecteur vitesse du flot (1). Donc les courbes intégrales (1) sont les caractéristiques de la forme $p\,dq - H dt$, c.q.f.d.

D. Théorème de l'invariant intégral de Poincaré-Cartan. Appliquons maintenant le lemme de Stokes. On obtient le théorème fondamental suivant.

Théorème. *Supposons que les courbes fermées γ_1 et γ_2 englobent un même tube de trajectoires (1). Alors les intégrales de la forme $p\,dq - H\,dt$ prennent la même valeur sur ces courbes:*

$$\oint_{\gamma_1} p\,dq - H\,dt = \oint_{\gamma_2} p\,dq - H\,dt.$$

La forme $p\,dq - H dt$ s'appelle *invariant intégral de Poincaré-Cartan* *).

D é m o n s t r a t i o n. Les trajectoires de phase sont les caractéristiques de la forme $p\,dq - H dt$; quant aux intégrales, elles prennent la même valeur sur les courbes fermées englobant un tube de caractéristiques d'après le lemme de Stokes, c.q.f.d.

---

*) En calcul variationnel $\int p\,dq - H\,dt$ s'appelle invariant intégral de Hilbert.

Considérons en particulier les courbes composées d'états simultanés, i.e. celles situées sur les plans $t = \text{const}$ (fig. 183). Le long de ces courbes $dt = 0$ et $\oint p\,dq - H dt = \oint p\,dq$. Du théorème précédent on déduit l'important

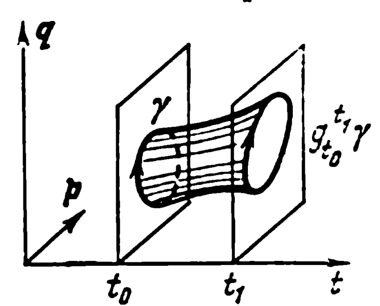


Fig. 183. Invariant intégral de Poincaré.

Corollaire 1. *Le flot conserve les intégrales de la forme* $p\,dq = p_1 dq_1 + \ldots + p_n dq_n$ *étendues aux courbes fermées.*

En effet, soit $g_{t_0}^{t_1}: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ une transformation de l'espace des phases $\{(p, q)\}$ réalisée par le flot entre la date t_0 et la date t_1 (i.e. $g_{t_0}^{t_1}(p_0, q_0)$ est solution des équations canoniques (1) avec les conditions initiales $p(t_0) = p_0$, $q(t_0) = q_0$). Soit γ une courbe quelconque fermée de l'espace $\mathbf{R}^{2n} \subset \mathbf{R}^{2n+1}$ ($t = t_0$). Alors $g_{t_0}^{t_1}\gamma$ est une courbe fermée dans l'espace $\mathbf{R}^{2n}$ ($t = t_1$) englobant le même tube des trajectoires de phase dans $\mathbf{R}^{2n+1}$. En vertu du théorème précédent, puisque $dt = 0$ sur γ et sur $g_{t_0}^{t_1}\gamma$, il vient $\oint\limits_{\gamma} p\,dq = \oint\limits_{g_{t_0}^{t_1}\gamma} p\,dq$, c.q.f.d.

La forme $p\,dq$ est appelée *invariant intégral relatif de Poincaré*. Il admet une interprétation géométrique simple. Soit σ une 2-chaîne orientée, $\gamma = \partial\sigma$. La formule de Stokes donne

$$\oint\limits_{\gamma} p\,dq = \iint\limits_{\sigma} dp \wedge dq.$$

On a démontré l'important

Corollaire 2. *Le flot conserve la somme des aires orientées des projections d'une surface sur les n plans arithmétiques* (p_i, q_i):

$$\iint\limits_{\sigma} dp \wedge dq = \iint\limits_{g_{t_0}^{t_1}\sigma} dp \wedge dq.$$

En d'autres termes la 2-forme $\omega^2 = dp \wedge dq$ est invariant intégral absolu du flot.

E x e m p l e. Pour $n = 1$ la forme ω^2 représente l'aire et nous obtenons le théorème de Liouville: le flot conserve l'aire.

E. Applications canoniques. Soit g une application différentiable de l'espace des phases $\mathbf{R}^{2n} = \{(p, q)\}$ dans $\mathbf{R}^{2n}$.

D é f i n i t i o n. L'application g est dite *canonique* si elle conserve la 2-forme $\omega^2 = \Sigma dp_i \wedge dq_i$.

Les raisonnements précédents montrent que cette définition peut prendre indifféremment l'une des trois formes:

1) $g^*\omega^2 = \omega^2$ (g conserve la 2-forme $\Sigma dp_i \wedge dq_i$);

2) $\iint\limits_{\sigma} \omega^2 = \iint\limits_{g\sigma} \omega^2 \ \forall \sigma$ (g conserve la somme des aires des projections d'une surface quelconque);

3) $\oint\limits_{\gamma} p\,dq = \oint\limits_{g\gamma} p\,dq$ (la forme $p\,dq$ est invariant intégral relatif de g).

E x e r c i c e. Montrer que les définitions 1) et 2) sont équivalentes à 3) dans le cas d'applications d'un domaine simplement connexe dans l'espace des phases $\mathbf{R}^{2n}$; dans le cas général $3 \Rightarrow 2 \Leftrightarrow 1$.

Les corollaires précédents peuvent encore s'énoncer:

Théorème. *Toute transformation de l'espace des phases par le flot est canonique.*

Soit $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ une transformation canonique: g conserve la forme ω^2. Alors g conserve également le carré extérieur de ω^2:

$$g^* (\omega^2 \wedge \omega^2) = \omega^2 \wedge \omega^2, \quad g^* (\omega^2)^k = (\omega^2)^k.$$

Les puissances extérieures de la forme $\Sigma dp_i \wedge dq_i$ sont proportionnelles aux formes

$$\omega^4 = \sum_{i<j} dp_i \wedge dp_j \wedge dq_i \wedge dq_j,$$

$$\omega^{2k} = \sum_{i_1 < \ldots < i_k} dp_{i_1} \wedge \ldots \wedge dp_{i_k} \wedge dq_{i_1} \wedge \ldots \wedge dq_{i_k}.$$

Ainsi on a démontré le

Théorème. *Les transformations canoniques conservent les invariants intégraux* $\omega^4, \ldots, \omega^{2n}$.

Géométriquement l'intégrale de la forme ω^{2k} sur une chaîne représente la somme des volumes orientés des projections de cette chaîne sur les plans arithmétiques $(p_{i_1}, \ldots, p_{i_k}, q_{i_1}, \ldots, q_{i_k})$.

En particulier, la forme ω^{2n} est proportionnelle à l'élément de volume et nous obtenons le

Corollaire. *Toute transformation canonique conserve l'élément de volume dans l'espace des phases*:

volume gD = volume D d'un domaine quelconque D.

En particulier, dans le cas d'un flot on obtient le

Corollaire. *Le flot* (1) *admet pour invariants intégraux les formes* $\omega^2, \omega^4, \ldots, \omega^{2n}$.

Le dernier de ces invariants est un volume, de sorte que nous avons de nouveau démontré le théorème de Liouville.

§ 45. Corollaires du théorème de l'invariant intégral de Poincaré-Cartan

Dans ce paragraphe on démontre que les transformations canoniques conservent la forme des équations de Hamilton, qu'une intégrale première des équations de Hamilton permet d'abaisser l'ordre du système de deux unités et que le mouvement d'un système naturel lagrangien se déroule sur une géodésique de l'espace de configuration muni d'une métrique riemannienne

A. Changement de variables dans les équations canoniques. L'invariance de la forme $p\,dq - H\,dt$ par les caractéristiques permet d'écrire les équations de mouvement d'un système quelconque de $2n+1$ coordonnées dans l'espace des phases élargi $\{(p, q, t)\}$.

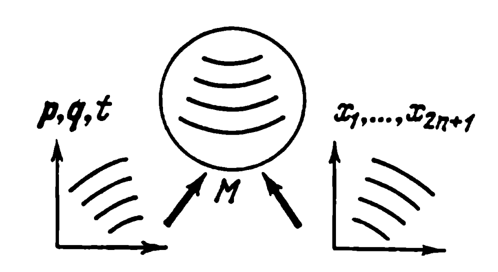


Fig. 184. Changement de variables dans les équations de Hamilton.

Soient $(x_1, \ldots, x_{2n+1})$ des fonctions coordonnées sur une carte de l'espace des phases élargi (que l'on suppose être une variété M, fig. 184). Les coordonnées (p, q, t) peuvent être considérées comme définissant une autre carte M. La forme $\omega^1 = p\,dq - H\,dt$ peut être envisagée comme une 1-forme différentielle sur M. A cette forme est intrinsèquement liée (i. e. ne dépend pas des cartes) une famille de lignes sur M : les caractéristiques. Sur la carte (p, q, t) ces lignes sont représentées par des trajectoires du flot

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q}, \quad \frac{dq}{dt} = \frac{\partial H}{\partial p} \tag{1}$$

de hamiltonien $H(p, q, t)$.

Supposons que dans les coordonnées $(x_1, \ldots, x_{2n+1})$ la forme ω^1 a pour expression

$$p\,dq - H\,dt = X_1\,dx_1 + \ldots + X_{2n+1}dx_{2n+1}$$

Théorème. *Sur la carte* (x_i) *les trajectoires* (1) *sont représentées par les caractéristiques de la forme* $\sum x_i dx_i$.

D é m o n s t r a t i o n. Les caractéristiques des formes $\sum X_i dx_i$ et $p\,dq - H\,dt$ représentent sur deux cartes différentes les

caractéristiques d'une même forme sur M. Or les courbes intégrales (1) sont les caractéristiques de la forme $p\,dq - H\,dt$. Donc leurs images sur la carte (x_i) sont les caractéristiques de la forme $\sum X_i dx_i$, c.q.f.d.

Corollaire. *Soit* $(P_1, \ldots, P_n;\ Q_1, \ldots, Q_n;\ T)$ *un système de coordonnées locales sur l'espace des phases élargi* $\{(p, q, t)\}$ *et* $K(P, Q, T)$, $S(P, Q, T)$ *des fonctions telles que*

$$p\,dq - H\,dt = P\,dQ - K\,dT + dS$$

(*le premier et le second membre sont des formes sur l'espace des phases élargi*).

Alors les trajectoires du flot (1) *sont représentées sur la carte* (P, Q, T) *par les courbes intégrales des équations canoniques*

$$\frac{dP}{dT} = -\frac{\partial K}{\partial Q}, \quad \frac{dQ}{dT} = \frac{\partial K}{\partial P}. \tag{2}$$

Démonstration. D'après le théorème précédent, les trajectoires (1) sont représentées par les caractéristiques de la forme $P\,dQ - K\,dt + dS$. Or dS n'exerce aucun effet sur les caractéristiques (puisque $d\,dS = 0$). Donc les images des trajectoires (1) sont les caractéristiques de la forme $P\,dQ - K\,dT$. En vertu du § 44, C, les caractéristiques d'une telle forme sont les courbes intégrales des équations canoniques (2), c.q.f.d.

En particulier soit $g\colon \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ une transformation canonique de l'espace des phases faisant correspondre au point de coordonnées (p, q) le point de coordonnées (P, Q).

Les fonctions $P(p, q)$ et $Q(p, q)$ peuvent être considérées comme de nouvelles coordonnées dans l'espace des phases.

Théorème. *Dans les nouvelles coordonnées* (P, Q) *les équations canoniques* (1) *se mettent sous la forme canonique* *)

$$\frac{dP}{dt} = -\frac{\partial K}{\partial Q}, \quad \frac{dQ}{dt} = \frac{\partial K}{\partial P} \tag{3}$$

d'ancienne fonction de Hamilton: $K(P, Q, t) = H(p, q, t)$.

Démonstration. Considérons la 1-forme $p\,dq - P\,dQ$ sur $\mathbf{R}^{2n}$. Pour toute courbe fermée γ on a (fig. 185)

$$\oint_\gamma p\,dq - P\,dQ = \oint_\gamma p\,dq - \oint_\gamma P\,dQ = 0$$

*) Dans certains ouvrages la propriété de conserver la forme canonique des équations de Hamilton est prise pour définition des transformations canoniques. En réalité cette définition n'est pas équivalente à celle généralement admise que nous avons citée plus haut. Par exemple, la transformation $P = 2p$, $Q = q$ conserve la forme hamiltonienne des équations de mouvement alors qu'elle n'est pas canonique selon notre définition.

en vertu de la canonicité de g. Donc $\int_{p_0, q_0}^{p_1, q_1} p\,dq - P\,dQ = S$ ne dépend pas du chemin d'intégration mais seulement de l'extrémité

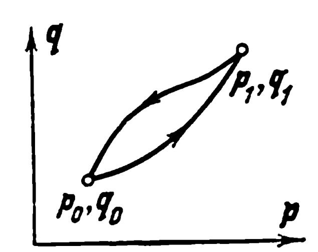


Fig. 185. Fermeture de la forme $p\,dq - P\,dQ$.

(p_1, q_1) (l'origine (p_0, q_0) étant fixe). Ainsi $dS = p\,dq - P\,dQ$. Dans l'espace des phases élargi on a donc

$$p\,dq - H dt = P\,dQ - H\,dt + dS,$$

et l'on peut appliquer le théorème précédent. On constate que (2) se transforme en (3), c.q.f.d.

Exercice. Soit $g(t): \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ une transformation canonique de l'espace des phases, fonction du paramètre t: $g(t)(p, q) = (P(p, q, t), Q(p, q, t))$. Montrer que dans les variables P, Q, t les équations canoniques (1) prennent la forme canonique avec une nouvelle fonction de Hamilton

$$K(P, Q, t) = H(p, q, t) - \frac{\partial S}{\partial t}, \quad \text{où } S(p_1, q_1, t) = \int_{p_0, q_0}^{p_1, q_1,} p\,dq - P\,dQ.$$

B. Abaissement de l'ordre par l'intégrale d'énergie. Supposons maintenant que le hamiltonien $H(p, q)$ ne dépende pas du temps. Les équations canoniques (1) admettront pour intégrale première $H(p(t), q(t)) = \text{const}$. Il s'avère qu'en se servant de cette intégrale on peut réduire la dimension $(2n + 1)$ de l'espace de deux unités et ramener le problème à l'intégration d'un système d'équations canoniques dans un espace de dimension $(2n - 1)$.

Supposons que (dans un certain domaine) l'équation $h = H(p_1, \ldots, p_n; q_1, \ldots, q_n)$ est résoluble en p_1:

$$p_1 = K(P, Q, T; h),$$

où $P = (p_2, \ldots, p_n)$; $Q = (q_2, \ldots, q_n)$; $T = -q_1$.

Il vient alors

$$p\,dq - H\,dt = P\,dQ - K\,dT - d(Ht) + t\,dH.$$

Soit γ une courbe intégrale des équations (1) située sur la surface $H(p, q) = h$ de dimension $2n$ de $\mathbf{R}^{2n+1}$. Alors γ est une caractéristique de la forme $p\,dq - H dt$ (fig. 186). Projetons l'espace des phases élargi $\mathbf{R}^{2n+1} = \{(p, q, t)\}$ sur l'espace des phases $\mathbf{R}^{2n} = \{(p, q)\}$. La

surface $H = h$ se projette sur une sous-variété M : $H(p, q) = h$ de $\mathbf{R}^{2n}$, et la courbe γ sur une courbe $\bar{\gamma}$ située sur cette sous-variété. Les grandeurs P, Q, T sont des coordonnées locales sur M.

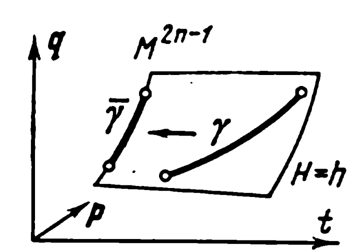


Fig. 186. Abaissement de l'ordre d'un système hamiltonien.

Exercice. Montrer que la *courbe* $\bar{\gamma}$ *est une caractéristique de la forme* $p\,dq = P\,dQ - K\,dT$ sur M.

Indication. $d(Ht)$ n'influe pas sur les caractéristiques, quant à dH elle est nulle sur M.

Or les caractéristiques de la forme $P\,dQ - K\,dT$ vérifient les équations de Hamilton (2). On a donc démontré le

Théorème. *Les trajectoires de phase des équations* (1) *vérifient sur la surface $H = h$ de M les équations canoniques*

$$\frac{dp_i}{dq_1} = \frac{\partial K}{\partial q_i}, \quad \frac{dq_i}{dq_1} = -\frac{\partial K}{\partial p_i} \qquad (i = 2, \ldots, n),$$

où la fonction $K(p_2, \ldots, p_n; q_2, \ldots, q_n : T, h)$ se définit à partir de l'équation $H(K, p_2, \ldots, p_n; -T, q_2, \ldots, q_n) = h$.

C. Principe de moindre action dans un espace des phases. Considérons dans l'espace des phases élargi $\{(p, q, t)\}$ la courbe intégrale γ des équations canoniques (1) joignant les points (p_0, q_0, t_0) et (p_1, q_1, t_1).

Théorème. *L'intégrale* $\int p\,dq - H\,dt$ *admet γ pour extrémale relativement aux variations de γ qui laissent les extrémités de cette courbe sur les sous-espaces* $(t = t_0, q = q_0)$ et $(t = t_1, q = q_1)$ *de dimension n.*

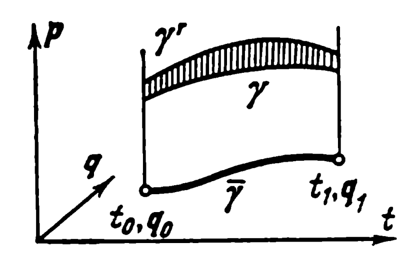


Fig. 187. Principe de moindre action dans l'espace des phases.

Démonstration. La courbe γ est une caractéristique de la forme $p\,dq - H\,dt$ (fig. 187). Donc l'intégrale $p\,dq - H\,dt$ sur un « parallélogramme infiniment petit passant par la direction du rotationnel » est nulle.

En d'autres termes, l'accroissement $\int\limits_{\gamma'} - \int\limits_{\gamma} p\,dq - H\,dt$ est un infiniment petit d'ordre supérieur par rapport à l'écart des courbes γ' et γ, c.q.f.d.

Si cette démonstration semble manquer de rigueur, on peut lui substituer le calcul :

$$\delta \int\limits_{\gamma} (p\dot{q} - H)\,dt = \int\limits_{\gamma} \left(\dot{q}\,\delta p + p\,\delta\dot{q} - \frac{\partial H}{\partial p}\delta p - \frac{\partial H}{\partial q}\delta q\right) dt =$$

$$= p\,\delta q\Big|_0^1 + \int\limits_{\gamma} \left[\left(\dot{q} - \frac{\partial H}{\partial p}\right)\delta p - \left(\dot{p} + \frac{\partial H}{\partial q}\right)\delta q\right] dt.$$

Nous constatons que les courbes intégrales des équations de Hamilton sont les seules extrémales de l'intégrale $\int p\,dq - H\,dt$ dans la classe des courbes γ dont les extrémités sont situées sur les sous-espaces $(t = t_0,\ q = q_0)$ et $(t = t_1,\ q = q_1)$ de l'espace des phases élargi. Le théorème est démontré.

Remarque. Le principe de moindre action sous la forme de Hamilton est un cas particulier de celui étudié plus haut. En effet, le long d'une extrémale on a

$$\int\limits_{t_0,\,q_0}^{t_1,\,q_1} p\,dq - H\,dt = \int\limits_{t_0}^{t_1} (p\dot{q} - H)\,dt = \int\limits_{t_0}^{t_1} L\,dt$$

(car le lagrangien L et le hamiltonien H se correspondent par la transformation de Legendre).

Soit d'autre part $\bar{\gamma}$ (fig. 188) la projection de l'extrémale γ sur le plan (q, t). A toute courbe $\bar{\gamma}'$ proche passant par les mêmes points (t_0, q_0) et (t_1, q_1) sur le

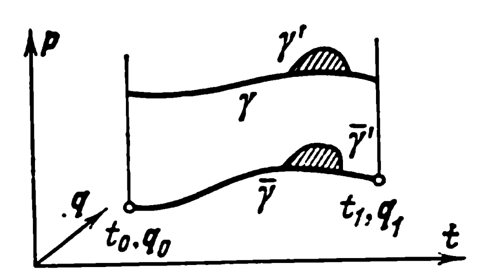


Fig. 188. Courbes de comparaison pour le principe de moindre action dans l'espace de configuration et dans l'espace des phases.

plan (q, t) associons une courbe γ' de l'espace des phases élargi $\{(p, q, t)\}$ en supposant que $p = \partial L/\partial\dot{q}$. Alors $\int\limits_{\gamma'} p\,dq - H\,dt = \int\limits_{\bar{\gamma}'} L\,dt$. Or d'après le théorème démontré $\delta \int\limits_{\gamma} p\,dq - H\,dt = 0$ quelles que soient les variations de la courbe γ (avec les conditions aux limites $(t = t_0,\ q = q_0)$ et $(t = t_1,\ q = q_1)$). En particulier ceci est vrai pour les variations de type spécial qui transforment γ en γ'. Donc $\bar{\gamma}$ est extrémale de $\int L\,dt$, c.q.f.d.

Dans le théorème démontré les courbes qui sont comparées avec γ appartiennent à une classe plus large que dans le principe de Hamilton : aucune restriction n'est imposée à la relation entre p et $\dot{q}$. Il peut paraître étonnant que ces deux principes n'en sont pas moins équivalents : l'extrémalité dans une classe plus étroite de variations ($p = \partial L/\partial \dot{q}$) entraîne l'extrémalité pour des variations quelconques. La raison en est que pour $\dot{q}$ fixe la quantité $p = \partial L/\partial \dot{q}$ extrémalise $p\dot{q} - H$ (cf. la définition de la transformation de Legendre § 14, page 68).

D. Principe de moindre action sous la forme de Maupertuis-Euler-Lagrange-Jacobi. Supposons maintenant que la fonction de Hamilton $H(p, q)$ ne dépende pas du temps. Alors $H(p, q)$ est intégrale première des équations de Hamilton (1). Projetons la surface $H(p, q) = h$ de l'espace des phases élargi $\{(p, q, t)\}$ sur l'espace $\{(p, q)\}$. On obtiendra la surface $H(p, q) = h$ de dimension $2n - 1$ dans $\mathbf{R}^{2n}$ que nous avons étudiée dans le point B et désignée par M.

Les trajectoires de phase des équations canoniques (1) issues de la surface M sont entièrement situées sur la surface M. Ce sont les caractéristiques de la forme $p\,dq = P\,dQ - K\,dT$ (dans les notations du point B) sur M. En vertu du théorème du point C, les courbes (1) sont extrémales du principe variationnel correspondant à cette forme. On a donc démontré le

Théorème. *Si la fonction de Hamilton $H = H(p, q)$ ne dépend pas du temps, alors les trajectoires de phase des équations canoniques (1) situées sur la surface M : $H(p, q) = h$ sont extrémales de l'intégrale $\int p\,dq$ dans la classe des courbes situées sur M et joignant les sous-espaces $q = q_0$ et $q = q_1$.*

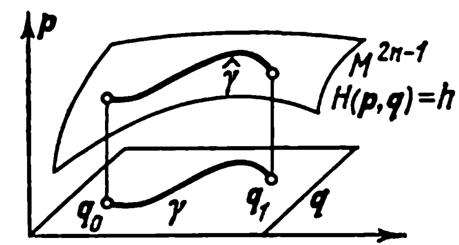


Fig. 189. Principe de Maupertuis.

Considérons maintenant la projection d'une extrémale de la surface M : $H(p, q) = h$ sur le plan q. Cette courbe joint les points q_0 et q_1.

Soit par ailleurs γ une autre courbe passant par les points q_0 et q_1 (fig. 189). Cette courbe γ est la projection d'une certaine courbe $\hat{\gamma}$ située sur la surface M. Plus exactement, choisissons sur γ un para-

mètre τ, $a \leqslant \tau \approx b$, $\gamma(a) = q_0$, $\gamma(b) = q_1$. Alors en chaque point q de la courbe γ est défini le vecteur vitesse $\dot{q} = \frac{d}{d\tau}\gamma(\tau)$ et l'impulsion $p = \partial L/\partial \dot{q}$ correspondante. Si le paramètre τ est choisi de telle sorte que $H(p, q) = h$, on obtient alors la courbe $\hat{\gamma}: q = \gamma(\tau)$, $p = \partial L/\partial \dot{q}$ sur la surface M. En appliquant le théorème précédent aux courbes $\hat{\gamma}$ sur M on obtient le

Corollaire. *Parmi toutes les courbes $q = \gamma(\tau)$ passant par deux points q_0 et q_1 du plan (q) et paramétrisées de telle sorte que la fonction de Hamilton prenne une valeur fixe $H(\partial L/\partial \dot{q}, q) = h$, la trajectoire des équations de la dynamique (1) est celle qui est extrémale de l'intégrale de l'« action réduite* :

$$\int_\gamma p\, dq = \int_\gamma p\dot{q}\, d\tau = \int_\gamma \frac{\partial L}{\partial \dot{q}}(\tau)\, \dot{q}(\tau)\, d\tau.$$

C'est le *principe de moindre action de Maupertuis* (*Euler-Lagrange-Jacobi*) *). Il est important de remarquer que l'intervalle $a \leqslant \tau \leqslant b$ qui paramétrise la courbe n'est pas fixe et peut varier suivant les courbes que l'on compare. Par contre, c'est l'énergie (la fonction de Hamilton) qui ne doit pas varier. Remarquons encore que le principe définit la forme de la trajectoire mais pas le temps : pour le déterminer il faut se servir de la constante d'énergie.

Le principe démontré prend une forme particulièrement simple si le système décrit un mouvement d'inertie sur une variété différentiable.

Théorème. *Un point matériel assujetti à rester sur une variété riemannienne différentiable se déplace sur une géodésique (i.e. sur une extrémale de la longueur $\int ds$).*

D é m o n s t r a t i o n. En effet, dans notre cas

$$H = L = T = \frac{1}{2}\left(\frac{ds}{d\tau}\right)^2, \quad \frac{\partial L}{\partial \dot{q}}\dot{q} = 2T = \left(\frac{ds}{d\tau}\right)^2.$$

Donc pour assurer à H une valeur fixe h le paramètre τ doit être choisi proportionnel à la longueur : $d\tau = \frac{ds}{\sqrt{2h}}$. L'intégrale de l'action réduite est alors égale à

$$\int_\gamma \frac{\partial L}{\partial \dot{q}}\dot{q}\, d\tau = \int_\gamma \sqrt{2h}\, ds = \sqrt{2h}\int_\gamma ds,$$

donc les extrémales sont géodésiques de notre variété, c.q.f.d.

*) « Dans presque tous les ouvrages, même dans les meilleurs, ce principe est présenté sous une forme telle qu'il est impossible de le comprendre. » (C. J a c o b i, *Cours de dynamique*, 1842-1843).

Dans le cas où l'énergie potentielle intervient, les trajectoires des équations de la dynamique sont également géodésiques d'une certaine métrique riemannienne.

Supposons que la métrique riemannienne ds^2 définisse l'énergie cinétique sur l'espace de configuration (de sorte que $T = \frac{1}{2}\left(\frac{ds}{dt}\right)^2$). Soit h une constante.

Théorème. *Définissons sur un domaine de l'espace de configuration où $U(q) < h$ une métrique riemannienne par la formule*

$$d\rho = \sqrt{h - U(q)}\, ds.$$

Alors les trajectoires du système d'énergie cinétique $T = \frac{1}{2}\left(\frac{ds}{dt}\right)^2$, d'énergie potentielle $U(q)$ et d'énergie totale h seront géodésiques dans la métrique $d\rho$.

D é m o n s t r a t i o n. En effet, dans notre cas $L = T - U$, $H = T + U$, $\frac{\partial L}{\partial \dot{q}} \dot{q} = 2T = \left(\frac{ds}{dt}\right)^2 = 2(h - U)$. Donc pour que H prenne une valeur fixe h le paramètre τ doit être choisi proportionnel à la longueur : $d\tau = \frac{ds}{\sqrt{2(h-U)}}$. L'intégrale de l'action réduite sera alors égale à

$$\int_\gamma \frac{\partial L}{\partial \dot{q}} \dot{q}\, d\tau = \int_\gamma \sqrt{2(h-U)}\, ds = \sqrt{2} \int_\gamma d\rho.$$

D'après le principe de Maupertuis les trajectoires sont géodésiques dans la métrique $d\rho$, c.q.f.d.

R e m a r q u e 1. La métrique $d\rho$ se déduit de ds par une homothétie de rapport $k > 1$ dépendant du point q mais pas de la direction. Donc les angles se confondent dans les métriques $d\rho$ et ds. Sur le bord du domaine $U \leqslant h$ la métrique $d\rho$ présente une singularité : plus on se rapproche du bord et plus la longueur ρ diminue. En particulier, la longueur de toute courbe située sur le bord même ($U = h$) est nulle.

R e m a r q u e 2. Si les extrémités d'une géodésique γ sont suffisamment proches l'une de l'autre, alors l'extremum de la longueur est un *minimum*. Ceci justifie l'expression « principe de moindre action ». L'examen des géodésiques sur la sphère unitaire (fig. 190) montre que dans le cas général l'extremum de l'action n'est pas forcément un minimum. Chaque arc de méridien est une géodésique, mais ne sont minimaux que les arcs qui sont inférieurs à π : l'arc $NS'M$ est plus court que l'arc de méridien NSM.

R e m a r q u e 3. Si h est supérieur au maximum U, sur l'espace de configuration, alors la métrique $d\rho$ ne présente pas de singularités. Et nous pouvons appliquer à l'étude des systèmes mécaniques

les théorèmes topologiques relatifs aux géodésiques sur des variétés riemanniennes.

Soit par exemple un tore T^2 sur lequel on a défini une métrique riemannienne. Parmi toutes les courbes fermées qui entourent T^2 m fois suivant un parallèle et n fois suivant un méridien il en existe

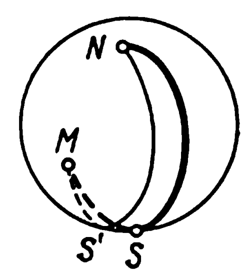


Fig. 190. Géodésique non minimale.

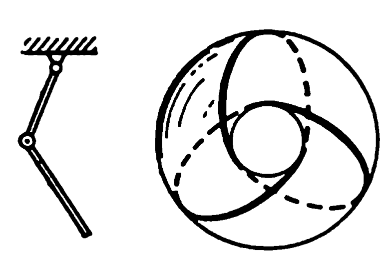


Fig. 191. Mouvement périodique du pendule double.

une qui est la plus courte (fig. 191). Cette courbe est une géodésique fermée (voir démonstration dans les cours de Calcul variationnel ou dans la « théorie de Morse »).

D'autre part, le tore T^2 est l'espace de configuration d'un pendule double plan. D'où il s'ensuit le

Théorème. *Quels que soient les entiers m, n, le pendule double admet un mouvement périodique tel qu'une masse effectue m révolutions et l'autre n pendant le même intervalle de temps.*

Bien plus, de tels mouvements périodiques existent pour une valeur quelconque suffisamment grande de la constante d'énergie h (h doit être supérieur à l'énergie potentielle dans la position supérieure).

Considérons encore l'exemple d'un solide ayant un point fixe dans un champ potentiel. L'espace de configuration SO(3) n'est pas simplement connexe: il contient des courbes non contractiles. Les raisonnements précédents entraînent le

Théorème. *Quel que soit le champ de forces potentiel, le solide admet au moins un mouvement périodique. Et de surcroît il existe des mouvements périodiques tels que la constante d'énergie h soit aussi grande que l'on veut.*

§ 46. Principe de Huygens

Les notions fondamentales de la mécanique hamiltonienne (les impulsions p, le hamiltonien H, la forme $p\,dq - H\,dt$, l'équation de Hamilton-Jacobi qui sera traitée plus bas) sont apparues lorsqu'on a appliqué aux principes variationnels généraux (et notamment au principe d'action stationnaire de Hamilton $\delta \int L\,dt = 0$) certaines notions naturelles et élémentaires de l'optique géométrique qui est régie par un principe variationnel particulier — le principe de Fermat.

A. Fronts d'ondes. Examinons brièvement *) les notions fondamentales de l'optique géométrique. En vertu du principe extrémal de Fermat, la *lumière traverse l'espace entre un point q_0 et un point q_1*

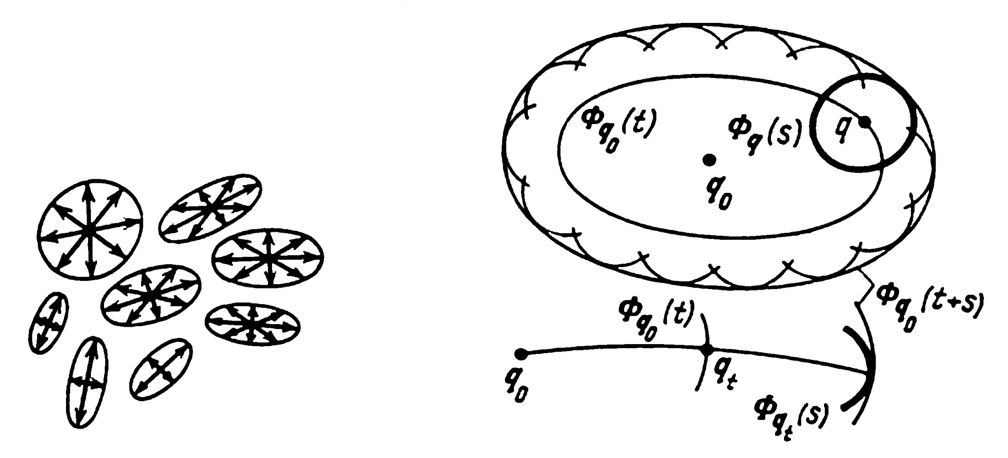


Fig. 192. Milieu non homogène anisotrope.

Fig. 193. Enveloppe de fronts d'ondes.

en un temps minimal. En outre la vitesse de la lumière peut dépendre aussi bien du point q (en milieu hétérogène) que de la direction du rayon (en milieu anisotrope, par exemple dans les cristaux).

On peut décrire les propriétés du milieu par la donnée dans l'espace tangent en chaque point q d'une surface (« indicatrice ») formée des extrémités de tous les vecteurs vitesse de propagation de la lumière au point q (fig. 192).

Supposons maintenant que $t > 0$. Considérons l'ensemble de tous les points q que la lumière peut atteindre au bout d'un temps inférieur ou égal à t en partant d'un point donné q_0. Le bord $\Phi_{q_0}(t)$ de cet ensemble est appelé *front d'onde* du point q_0 au bout du temps t ; il est composé des points que la lumière peut atteindre au bout du temps t et pas plus vite.

Entre les fronts d'ondes correspondant à divers t il existe une relation remarquable découverte par Huygens (fig. 193).

Théorème de Huygens. *Considérons le front d'onde $\Phi_{q_0}(t)$ d'un point q_0 au bout d'un temps t. Construisons pour chaque point q de ce front le front d'onde Φ_q (s) au bout d'un temps s. Alors le front d'onde $\Phi_{q_0}(s+t)$ du point q_0 au bout du temps $s+t$ sera l'enveloppe des fronts Φ_q (s), $q \in \Phi_{q_0}(t)$.*

En effet, soit $q_{t+s} \in \Phi_{q_0}(t+s)$. Il existe alors un chemin entre q_0 et q_{t+s} que la lumière parcourt en un temps $t+s$ et il n'en est pas de plus court. Considérons sur ce chemin le point q_t que la

*) Nous ne rechercherons pas la rigueur ici, nous supposerons que tous les déterminants sont non nuls, etc. Les démonstrations des théorèmes ultérieurs ne souffriront pas des raisonnements semi-euristiques de ce point.

lumière atteint au bout du temps t. Il n'existe pas de plus court chemin pour aller de q_0 à q_t, sinon le chemin q_0q_{t+s} ne serait pas le plus court. Donc le point q_t est situé sur le front $\Phi_{q_0}(t)$. De la même façon le chemin q_tq_{t+s} est parcouru dans le temps s par la lumière et ce chemin est le plus court. Donc le point q_{t+s} est situé sur le front

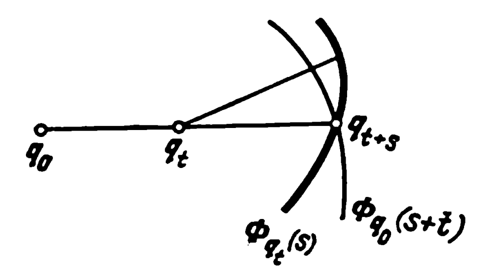


Fig. 194. Démonstration du théorème de Huygens.

$\Phi_{q_t}(s)$ du point q_t au bout du temps s. Montrons que les fronts $\Phi_{q_t}(s)$ et $\Phi_{q_0}(t+s)$ sont tangents au point q_{t+s}. En effet, s'ils se coupaient (fig. 194), on pourrait atteindre certains points de $\Phi_{q_0}(t+s)$ en un temps inférieur à s en partant de q_t et inférieur à $s+t$ en partant de q_0. Ceci contredit la définition de $\Phi_{q_0}(t+s)$; donc les fronts $\Phi_{q_t}(s)$ et $\Phi_{q_0}(t+s)$ sont tangents au point q_{t+s}, c.q.f.d.

Le théorème que nous venons de démontrer traduit le *principe de Huygens*. On aurait pu visiblement remplacer le point q_0 par une courbe, une surface ou d'une façon générale par un ensemble fermé, l'espace $\{q\}$ de dimension trois par une variété différentiable quelconque et la propagation de la lumière par la propagation d'une perturbation quelconque à caractère « local ».

Le principe de Huygens conduit à deux descriptions de la propagation. D'abord nous pouvons observer les *rayons*, i.e. les plus courts chemins de propagation de la lumière. Dans ce cas le caractère local de propagation est défini par le vecteur vitesse $\dot{q}$. Si la direction du rayon est connue, le module du vecteur vitesse est défini par les propriétés du milieu (par l'indicatrice).

Ensuite nous pouvons observer le front d'onde.

Supposons que dans l'espace $\{q\}$ est donnée une métrique riemannienne. On peut alors parler de *vitesse du front d'onde*. Considérons par exemple la propagation de la lumière dans un milieu remplissant l'espace euclidien ordinaire. Le mouvement du front d'onde peut être caractérisé par un vecteur p orthogonal au front que l'on construit de la manière suivante.

Pour chaque point q_0 définissons la fonction $S_{q_0}(q)$ comme la *longueur optique du chemin entre* q_0 *et* q, i.e. le temps minimal de propagation de la lumière de q_0 à q.

Dans l'ensemble de niveau $\{q: S_{q_0}(q) = t\}$ on reconnaît le front d'onde $\Phi_{q_0}(t)$ (fig. 195). Le gradient de la fonction S (au sens de la métrique définie plus haut) est perpendiculaire au front

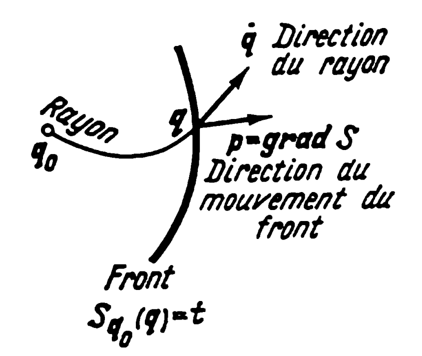


Fig. 195. Directions du rayon et du front d'ondes.

d'onde et caractérise son mouvement. En outre, plus le gradient est *élevé*, plus le mouvement du front est *lent*. C'est pourquoi Hamilton a appelé le vecteur

$$p = \frac{\partial S}{\partial q}$$

vecteur de lenteur normale du front.

Les directions du rayon $\dot{q}$ et du mouvement du front ne sont pas les mêmes dans un milieu anisotrope. Cependant elles sont liées par une relation simple que l'on déduit aisément du principe de

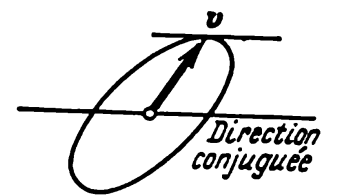


Fig. 196. Hyperplan conjugué.

Huygens. On rappelle que les propriétés du milieu sont caractérisées en chaque point par la surface des vecteurs vitesse de la lumière: l'indicatrice.

Définition. La direction de l'hyperplan tangent à l'indicatrice en un point v est conjuguée de la direction de v (fig. 196).

Théorème. *La direction du front d'onde $\Phi_{q_0}(t)$ au point q_t est conjuguée de la direction du rayon $\dot{q}$.*

Démonstration. Considérons un point q_τ (fig. 197) du rayon q_0q_t, $0 \leqslant \tau \leqslant t$. Soit ε très petit. Alors l'écart est de l'ordre de $O(\varepsilon^2)$ entre le front $\Phi_{q_{t-\varepsilon}}(\varepsilon)$ et la transformée de l'indicatrice du point q_t par une homothétie de rapport ε. D'après le principe de Huygens, le front $\Phi_{q_{t-\tau}}(\varepsilon)$ est tangent en q_t au front $\Phi_{q_0}(t)$. En passant à la limite pour $\varepsilon \to 0$ on obtient le théorème énoncé.

Si l'on change la métrique auxiliaire introduite pour définir le vecteur p, alors variera la notion de vitesse du front et partant le module et la direction du vecteur p. Cependant la forme différentielle $p\,dq = dS$ sur l'espace $\{q\} = \mathbf{R}^3$ a été définie indépendam-

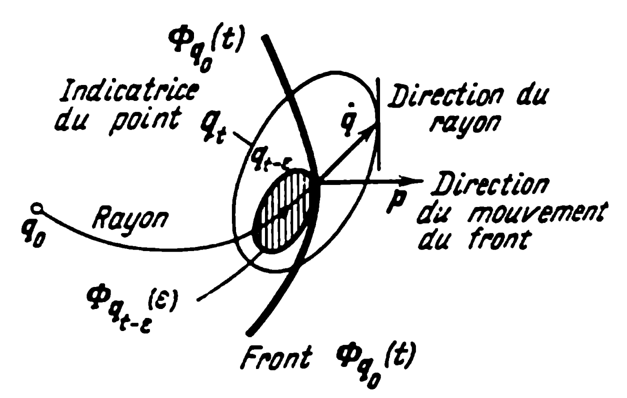


Fig. 197. Conjugaison des directions du rayon et du front.

ment de cette métrique; sa valeur ne dépend que du front choisi (ou du rayon). Sur l'hyperplan conjugué du vecteur vitesse de la lumière, cette forme est nulle et sur le vecteur vitesse *) elle est égale à l'unité.

B. Analogie entre l'optique et la mécanique. Revenons maintenant à la mécanique. Ici aussi les trajectoires du mouvement sont les extrémales d'un principe variationnel et l'on peut bâtir la mécanique à l'image de l'optique géométrique dans un espace multidimensionnel. C'est précisément ainsi qu'a procédé Hamilton; nous nous bornerons à citer les notions optiques qui ont conduit Hamilton aux principales notions de mécanique.

Optique	Mécanique
Milieu optique	Espace de configuration élargi $\{(q, t)\}$
Principe de Fermat	Principe de Hamilton $\delta \int L\,dt = 0$
Rayons	Trajectoires $q(t)$
Indicatrice	Lagrangien L
Lenteur normale du front p	Impulsion p
Expression de p en fonction de la vitesse $\dot{q}$ du rayon	Transformation de Legendre
1-forme $p\,dq$	1-forme $p\,dq - H\,dt$

*) Donc les vecteurs p correspondant à tous les fronts passant par un point donné ne sont pas arbitraires mais assujettis à une condition : les valeurs admissibles de p couvrent dans l'espace $\{p\}$ une hypersurface duale de l'indicatrice des vitesses.

Seuls la longueur du chemin $S_{q_0}(q)$ et le principe de Huygens n'ont pas été cités. Leurs analogues mécaniques sont l'*action* et l'*équation de Hamilton — Jacobi* que nous allons traiter maintenant.

C. L'action comme une fonction de q et t.

D é f i n i t i o n. On appelle *action* $S(q, t)$ l'intégrale

$$S_{q_0,t_0}(q, t) = \int_\gamma L\, dt$$

le long d'une extrémale γ passant par les points (q_0, t_0) et (q, t).

Pour que cette définition soit correcte quelques précautions sont à prendre: il faut exiger que les extrémales issues du point (q_0, t_0) ne se coupent plus ailleurs et forment un champ central

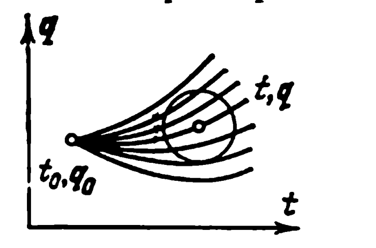


Fig. 198. Champ central d'extrémales.

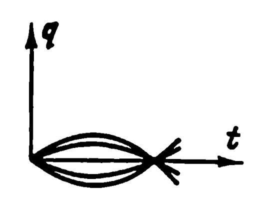


Fig. 199. Extrémale à point focal ne pouvant être incluse dans un champ central.

d'extrémales (fig. 198). D'une façon plus précise, faisons correspondre à chaque couple $(\dot{q}_0, t)$ le point (q, t), extrémité de l'extrémale vérifiant les conditions initiales $q(0) = q_0$, $\dot{q}(0) = \dot{q}_0$. On dit que l'*extrémale* γ *est incluse dans un champ central* si l'application $(\dot{q}_0, t) \to (q, t)$ est régulière (au point correspondant à l'extrémale considérée, et partant dans un de ses voisinages).

On démontre que pour $|t - t_0|$ suffisamment petit l'extrémale γ fait partie d'un champ central *).

Soit maintenant un voisinage suffisamment petit de l'extrémité (q, t) de notre extrémale γ. Chaque point de ce voisinage est relié au point (q_0, t_0) par une seule extrémale du champ central considéré. Cette extrémale dépend différentiablement du point final (q, t). Donc l'action

$$S_{q_0 t_0}(q, t) = \int_\gamma L\, dt$$

est correctement définie dans le voisinage indiqué.

*) E x e r c i c e. Montrer que ce n'est pas le cas pour $t - t_0$ grand.
I n d i c a t i o n. $\ddot{q} = -q$ (fig. 199).

En géométrique optique nous avons étudié la différentielle de la longueur optique d'un chemin. Il est naturel d'envisager ici la différentielle de l'action.

Théorème. *La différentielle de l'action (si le point initial est fixe) est égale à*

$$dS = p\,dq - H\,dt,$$

où $p = \partial L/\partial \dot{q}$ et $H = p\dot{q} - L$ dépendent de la vitesse finale $\dot{q}$ de la trajectoire γ.

D é m o n s t r a t i o n. Relevons chaque extrémale de l'espace $\{(q, t)\}$ dans l'espace des phases élargi $\{(p, q, t)\}$ en posant $p = \partial L/\partial \dot{q}$, i. e. en remplaçant l'extrémale par une trajectoire de phase. Nous obtenons alors dans l'espace des phases élargi une variété

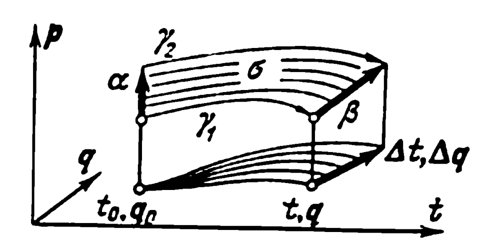


Fig. 200. Calcul de la différentielle de la fonction d'action.

de dimension $n + 1$ composée des trajectoires de phase, i.e. des caractéristiques de la forme $p\,dq - H\,dt$. Donnons maintenant au point final (q, t) l'accroissement $(\Delta q, \Delta t)$ et considérons la famille d'extrémales qui joignent le point (q_0, t_0) aux points du segment $q + \theta\,\Delta q$, $t + \theta\,\Delta t$, $0 \leqslant \theta \leqslant 1$ (fig. 200). Dans l'espace des phases nous obtenons un quadrilatère σ constitué de caractéristiques de la forme $p\,dq - H\,dt$, dont le bord

$$\partial\sigma = \gamma_1 - \gamma_2 + \beta - \alpha$$

est composé de deux trajectoires de phase γ_1 et γ_2, d'un segment de courbe α contenue dans l'espace $\{(q = q_0,\ t = t_0)\}$ et d'un segment de courbe β dont la projection est le segment $(\Delta q, \Delta t)$. Comme σ est composé de caractéristiques de la forme $p\,dq - H\,dt$ on a

$$0 = \iint_{\sigma} d(p\,dq - H\,dt) = \int_{\partial\sigma} p\,dq - H\,dt =$$
$$= \int_{\gamma_1} - \int_{\gamma_2} + \int_{\beta} - \int_{\alpha} p\,dq - H\,dt.$$

Or sur le segment α on a $dq = 0$, $dt = 0$. Sur les trajectoires de phase γ_1 et γ_2 on a $p\,dq - H\,dt = L\,dt$ (§ 45, C). Donc la différence $\int_{\gamma_2} - \int_{\gamma_1} p\,dq - H\,dt$ est égale à l'accroissement de l'action et

il vient

$$\int_\beta p\,dq - H\,dt = S(q+\Delta q,\ t+\Delta t) - S(q,\ t).$$

Si maintenant $\Delta q \to 0$, $\Delta t \to 0$, alors

$$\int_\beta p\,dq - H\,dt = p\,\Delta q - H\,\Delta t + o(\Delta t,\ \Delta q),$$

c.q.f.d.

Nous voyons que la forme $p\,dq - H\,dt$, artificiellement introduite auparavant, apparaît d'elle-même lorsqu'on établit une analogie entre l'action et la longueur optique du chemin.

D. Equation de Hamilton-Jacobi. Rappelons que le vecteur de lenteur normale p ne peut pas être arbitraire: il est assujetti à la condition $p\dot{q} = 1$ qui dérive du principe de Huygens. Une condition analogue est imposée au gradient de l'action S.

Théorème. *L'action vérifie l'équation*

$$\frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial q},\ q,\ t\right) = 0. \qquad (1)$$

Cette équation non linéaire du premier ordre aux dérivées partielles s'appelle *équation de Hamilton-Jacobi.*

Pour démontrer le théorème il suffit de remarquer que d'après le théorème précédent

$$\frac{\partial S}{\partial t} = -H(p,\ q,\ t),\quad p = \frac{\partial S}{\partial q}.$$

La relation que nous avons établie entre les trajectoires du système mécanique (« rayons ») et l'équation aux dérivées partielles (« fronts d'ondes ») est d'un double usage.

D'abord, certaines solutions de l'équation (1) peuvent être utilisées pour l'intégration d'équations différentielles ordinaires de la dynamique. C'est d'ailleurs en cela que consiste la méthode de Jacobi d'intégration des équations canoniques de Hamilton qui fera l'objet du paragraphe suivant.

Ensuite, la relation établie permet de ramener l'intégration de l'équation aux dérivées partielles (1) à l'intégration d'un système d'équations différentielles ordinaires de Hamilton.

Penchons-nous plus en détail sur ce point. Soit à résoudre le problème de Cauchy pour les équations de Hamilton-Jacobi (1)

$$S(q,\ t_0) = S_0(q),\quad \frac{\partial S}{\partial t} + H\left(\frac{\partial S}{\partial q},\ q,\ t\right) = 0. \qquad (2)$$

Pour cela considérons le système d'équations canoniques de Hamilton

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}.$$

Soient les conditions initiales (fig. 201) :

$$q(t_0) = q_0, \quad p(t_0) = \left.\frac{\partial S_0}{\partial q}\right|_{q_0}.$$

La solution qui vérifie ces conditions initiales est représentée dans l'espace $\{(q, t)\}$ par une courbe $q = q(t)$, extrémale du principe $\delta \int L\, dt = 0$ (où le lagrangien $L(q, \dot{q}, t)$ est la transformée de Legen-

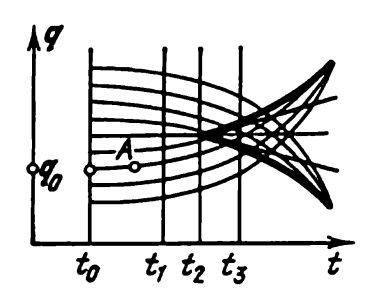


Fig. 201. Caractéristiques du problème de Cauchy pour l'équation de Hamilton-Jacobi.

Fig. 202. Fonction d'action comme solution de l'équation de Hamilton-Jacobi.

dre du hamiltonien $H(p, q, t)$ par rapport à p). Cette extrémale s'appelle *caractéristique* du problème (2) issue du point q_0.

Si une valeur t_1 est suffisamment proche de t_0, alors les caractéristiques issues de points proches de q_0 ne se coupent pas pour $t_0 \leqslant t \leqslant t_1$, $|q - q_0| < R$. Par ailleurs, les valeurs de q_0 et t peuvent être prises pour coordonnées d'un point A dans le domaine $|q| < R$, $t_0 \leqslant t \leqslant t_1$ (fig. 201).

Construisons maintenant l'« action avec la condition initiale S_0 » :

$$S(A) = S_0(q_0) + \int_{q_0, t_0}^{A} L(q, \dot{q}, t)\, dt \tag{3}$$

(l'intégration s'effectue le long d'une caractéristique menant au point A).

Théorème. *La fonction* (3) *est solution du problème* (2).

En effet, la condition initiale est visiblement réalisée. On s'assure que l'équation de Hamilton-Jacobi est vérifiée comme dans le théorème relatif à la différentielle de l'action (fig. 202).

D'après le lemme de Stokes $\int\limits_{\gamma_1} - \int\limits_{\gamma_2} + \int\limits_{\beta} - \int\limits_{\alpha} p\,dq - H\,dt = 0$. Or sur α on a $dt = 0$, $p = \partial S_0/\partial q$. Donc

$$\int\limits_{\alpha} p\,dq - H\,dt = \int\limits_{\alpha} p\,dq = \int\limits_{\alpha} dS_0 = S_0(q_0 + \Delta q) - S_0(q_0).$$

Puis, γ_1 et γ_2 étant des trajectoires de phase, il vient

$$\int\limits_{\gamma_1, \gamma_2} p\,dq - H\,dt = \int\limits_{\gamma_1, \gamma_2} L\,dt.$$

Donc

$$\int\limits_{\beta} p\,dq - H\,dt = \left[S_0(q_0 + \Delta q) + \int\limits_{\gamma_2} L\,dt\right] - \left[S_0(q_0) + \int\limits_{\gamma_1} L\,dt\right] = S(A + \Delta A) - S(A).$$

Lorsque $\Delta t \to 0$, on obtient $\frac{\partial S}{\partial t} = -H$, $\frac{\partial S}{\partial q} = p$, ce qui démontre le théorème.

E x e r c i c e. Démontrer l'unicité de la solution du problème (2).

I n d i c a t i o n. Dériver S sur des caractéristiques.

E x e r c i c e. Résoudre le problème de Cauchy (2) pour $H = p^2/2$, $S_0 = q^2/2$.

E x e r c i c e. Tracer les graphes des « fonctions » multivoques $S(q)$ et $p(q)$ pour $t = t_3$ (fig. 201).

Réponse. Voir fig. 203.

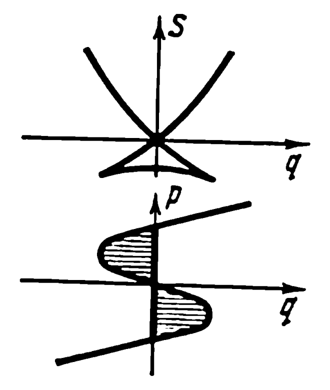


Fig. 203. Singularité type d'une solution de l'équation de Hamilton-Jacobi.

Au point double du graphe de S correspond sur le graphe de p la *droite de Maxwell*: les aires hachurées sont égales. Le graphe de $S(q, t)$ présente au point (q_0, t_2) une singularité appelée *queue d'aronde*.

§ 47. Méthode de Jacobi-Hamilton d'intégration des équations canoniques de Hamilton

Dans ce paragraphe on donne la définition de la fonction génératrice de la transformation canonique libre.

L'idée de la méthode de Jacobi-Hamilton est la suivante: la forme canonique des équations de mouvement ainsi que la fonction

de Hamilton (§ 45, A) sont invariantes par un changement canonique de coordonnées.

Si donc l'on réussit à trouver une transformation canonique qui fasse prendre à la fonction de Hamilton une forme telle que les équations canoniques soient intégrables, il sera alors possible d'intégrer les équations canoniques initiales. La construction d'une telle transformation canonique se ramène à la recherche d'un assez grand nombre de solutions de l'équation aux dérivées partielles de Hamilton-Jacobi. Cette équation doit être vérifiée par la fonction génératrice de la transformation canonique cherchée.

Avant de passer à l'appareil des fonctions génératrices, remarquons que malheureusement il n'est pas invariant et qu'il fait largement appel à la structure arithmétique de l'espace des phases $\{(p, q)\}$. En conséquence il faudra se servir des dérivées partielles, or leur notation même est déjà entachée d'ambiguïté *).

A. Fonction génératrice. Soient $2n$ fonctions $P(p, q)$, $Q(p, q)$ de $2n$ variables p, q définissant une transformation canonique $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$. Alors la 1-forme: $p\,dq - P\,dQ$ est une différentielle totale (§ 45, A):

$$p\,dq - P\,dQ = dS(p, q). \tag{1}$$

E x e r c i c e. Montrer que la réciproque est vraie, i.e. si cette forme est une différentielle totale, la transformation est canonique.

Supposons maintenant qu'au voisinage d'un point (p_0, q_0) on puisse prendre pour coordonnées indépendantes (Q, q). En d'autres termes, *supposons qu'est différent de zéro au point* (p_0, q_0) *le jacobien*

$$\det \frac{\partial (Q, q)}{\partial (p, q)} = \det \frac{\partial Q}{\partial p} \neq 0.$$

De telles transformations canoniques sont dites *libres*. Notamment la fonction S peut être exprimée localement en fonction de ces coordonnées:

$$S(p, q) = S_1(Q, q).$$

D é f i n i t i o n. La fonction $S_1(Q, q)$ est appelée *fonction génératrice* de la transformation canonique g.

Soulignons que la fonction S_1 n'est pas une fonction définie sur l'espace des phases $\mathbf{R}^{2n}$: elle est en effet définie sur le produit

*) Il importe de comprendre clairement que sur le plan (x, y) la quantité $\partial u/\partial x$ ne dépend pas seulement de la fonction qui a été prise pour x, mais aussi de celle qui a été prise pour y: dans les nouvelles variables (x, z) la valeur de $\partial u/\partial x$ sera différente. Il aurait fallu écrire

$$\left.\frac{\partial u}{\partial x}\right|_{y=\text{const}}, \quad \left.\frac{\partial u}{\partial x}\right|_{z=\text{const}}.$$

cartésien $\mathbf{R}_q^n \times \mathbf{R}_Q^n$ de deux espaces arithmétiques de dimension n dont les points sont désignés par q et Q.

De (1) il s'ensuit que les « dérivées partielles » de S_1 sont

$$\frac{\partial S_1(Q, q)}{\partial q} = p, \quad \frac{\partial S_1(Q, q)}{\partial Q} = -P. \tag{2}$$

Il s'avère qu'inversement toute fonction S_1 définit une transformation canonique g d'après les formules (2).

Théorème. *Soit $S_1(Q, q)$ une fonction définie au voisinage d'un point (Q_0, q_0) du produit cartésien de deux espaces arithmétiques euclidiens de dimension n. Si*

$$\det \left.\frac{\partial^2 S_1}{\partial Q\, \partial q}\right|_{Q_0, q_0} \neq 0,$$

alors la fonction S_1 est la génératrice d'une transformation canonique libre.

Démonstration. Considérons l'équation en Q

$$\frac{\partial S_1(Q, q)}{\partial q} = p.$$

D'après le théorème des fonctions implicites, cette équation est résoluble et définit au voisinage du point $\left(q_0, p_0 = \left.\frac{\partial S_1(Q,q)}{\partial q}\right|_{Q_0, q_0}\right)$ une fonction $Q(p, q)$ telle que $Q(p_0, q_0) = Q_0$. En effet, le jacobien est justement égal à $\det \left.\frac{\partial^2 S_1(Q, q)}{\partial Q\, \partial q}\right|_{Q_0, q_0}$, or par hypothèse il est nul.

Soit maintenant la fonction

$$P_1(Q, q) = -\frac{\partial}{\partial Q} S_1(Q, q).$$

Posons

$$P(p, q) = P_1(Q(p, q), q).$$

Alors la transformation locale $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ qui enverra (p, q) en $P(p, q)$, $Q(p, q)$ sera canonique et admettra S_1 pour fonction génératrice puisque par construction

$$p\, dq - P\, dQ = \frac{\partial S_1(Q, q)}{\partial q}\, dq + \frac{\partial S_1(Q, q)}{\partial Q}\, dQ.$$

Elle est libre puisque $\det \frac{\partial Q}{\partial p} = \det \left(\frac{\partial^2 S_1(Q, q)}{\partial Q\, \partial q}\right)^{-1} \neq 0$. Le théorème est démontré.

La transformation $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ est donnée en général par $2n$ fonctions de $2n$ variables. Nous remarquons qu'elle est définie par une *seule* fonction de $2n$ variables: sa fonction génératrice. On conçoit aisément l'avantage des fonctions génératrices dans les calculs faisant intervenir les transformations canoniques, un avan-

tage qui est d'autant plus grand que le nombre des variables est élevé.

B. Equation de Hamilton-Jacobi pour la fonction génératrice. Remarquons que les équations canoniques dans lesquelles le hamiltonien H dépend uniquement des coordonnées Q s'intègrent facilement. En effet, si $H = K(Q)$, les équations canoniques s'écrivent

$$\dot{Q} = 0, \quad \dot{P} = \frac{\partial K}{\partial Q}, \tag{3}$$

d'où l'on déduit immédiatement

$$Q(t) = Q(0), \quad P(t) = P(0) + t\frac{\partial K}{\partial Q}\Big|_{Q_0}.$$

Cherchons maintenant les transformations canoniques qui ramènent le hamiltonien $H(p, q)$ à la forme $K(Q)$. Cherchons pour cela la fonction génératrice $S(Q, q)$ d'une telle transformation. De (2) nous obtenons la condition

$$H\left(\frac{\partial S(Q, q)}{\partial q}, q, t\right) = K(Q), \tag{4}$$

où *après* dérivation il faudra remplacer q par $q(P, Q)$. Remarquons que pour Q fixe l'équation (4) prend la forme de l'équation de Hamilton-Jacobi.

Théorème de Jacobi. *Si l'on a trouvé une solution $S(Q, q)$ de l'équation de Hamilton-Jacobi* (4) *dépendant de n paramètres Q_i* *) *et telle que* $\det \frac{\partial^2 S}{\partial Q\, \partial q} \neq 0$, *alors les équations canoniques*

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p} \tag{5}$$

s'intègrent explicitement par quadratures. Les fonctions $Q(p, q)$ définies par les équations $\frac{\partial S(Q, q)}{\partial q} = p$ sont n intégrales premières des équations (5).

D é m o n s t r a t i o n. Considérons la transformation canonique de fonction génératrice $S(Q, q)$. En vertu de (2) on a $p = \frac{\partial S}{\partial q}(Q, q)$, d'où l'on déduit $Q(p, q)$. Calculons la fonction $H(p, q)$ dans les nouvelles coordonnées P, Q. On a $H(p, q) = H\left(\frac{\partial S}{\partial q}(Q, q), q\right)$. Pour trouver le hamiltonien dans les nouvelles coordonnées il aurait fallu dans cette relation remplacer (après dérivation) q par son expression en fonction de P et Q. Or en vertu de (4) l'expression de cette fonction ne dépend pas de q, de sorte que l'on a tout simple-

*) On appelle *intégrale complète* d'une équation une famille à n paramètres de solutions de (4).

ment

$$H(p, q) = K(Q).$$

Donc, dans les nouvelles variables, les équations (5) s'écrivent sous la forme (3) d'où résulte immédiatement le théorème de Jacobi.

Le théorème de Jacobi ramène la résolution du système d'équations différentielles (5) à la recherche de l'intégrale complète de l'équation aux dérivées partielles (4). Il pourrait paraître étonnant que cette réduction d'un problème simple à un problème plus compliqué constitue une méthode efficace de résolution. Et pourtant nous verrons que c'est l'une des plus puissantes méthodes d'intégration analytique et de nombreux problèmes résolus par Jacobi sont inaccessibles aux autres méthodes.

C. Exemples. Considérons l'attraction de deux centres fixes. L'étude du mouvement des satellites artificiels a suscité un intérêt accru pour ce problème. Il est clair que deux centres attractifs

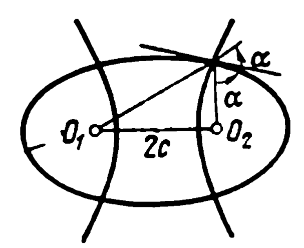


Fig. 204. Coordonnées elliptiques.

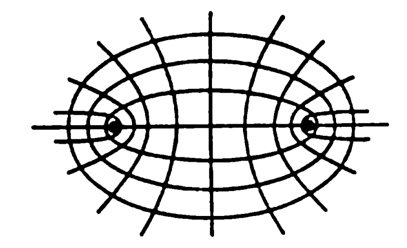


Fig. 205. Ellipses et hyperboles homofocales.

proches situés sur un axe z approximisent l'attraction d'un ellipsoïde légèrement allongé le long de cet axe. Malheureusement la Terre n'est pas allongée, mais au contraire aplatie. La solution consiste à placer les centres en des points imaginaires de coordonnées $\pm i\varepsilon$ sur l'axe z. Les formules analytiques de la solution restent de toute évidence en vigueur dans le domaine complexe. On obtient donc une approximation du champ d'attraction de la Terre, une approximation dans laquelle les équations de mouvement s'intègrent exactement et qui est plus proche de la réalité que l'approximation de Kepler (qui considère la Terre comme un point).

Considérons pour simplifier l'attraction de deux masses fixes égales dans le plan. La méthode de Jacobi fait appel à un système de coordonnées convenables dites coordonnées elliptiques. Soient deux points fixes O_1 et O_2, $2c$ la distance de ces points (fig. 204), r_1 et r_2 les distances de ces points à une masse en mouvement. Les coordonnées elliptiques ξ, η se définissent comme la somme et la différence des distances r_1 et r_2: $\xi = r_1 + r_2$, $\eta = r_1 - r_2$.

Exercice. Exprimer la fonction de Hamilton dans les coordonnées elliptiques.

Solution. Les lignes ξ = const sont des ellipses de foyers O_1 et O_2, les lignes η = const des hyperboles de mêmes foyers (fig. 205). Ellipses et hyperboles sont mutuellement orthogonales, donc $ds^2 = a^2 d\xi^2 + b^2 d\eta^2$. Trouvons les coefficients a et b. Lorsque le mouvement s'effectue sur une ellipse on a $dr_1 = ds\cos\alpha$, $dr_2 = -ds\cos\alpha$, $d\eta = 2\cos\alpha\, ds$. Sur une hyperbole on a $dr_1 = ds\sin\alpha$, $dr_2 = ds\sin\alpha$, $d\xi = 2\sin\alpha\, ds$. Donc $a = (2\sin\alpha)^{-1}$, $b = (2\cos\alpha)^{-1}$.

D'autre part, du triangle O_1MO_2 on déduit

$$r_1^2 + r_2^2 + 2r_1r_2\cos 2\alpha = 4c^2,$$

d'où

$$\cos^2\alpha - \sin^2\alpha = \frac{4c^2 - r_1^2 - r_2^2}{2r_1r_2},$$

$$\cos^2\alpha + \sin^2\alpha = \frac{2r_1r_2}{2r_1r_2},$$

$$\cos^2\alpha = \frac{4c^2 - (r_1 - r_2)^2}{4r_1r_2}, \quad \sin^2\alpha = \frac{(r_1 + r_2)^2 - 4c^2}{4r_1r_2}.$$

Mais si $ds^2 = \sum a_i^2 dq_i^2$, alors $T = \sum a_i^2 \frac{\dot{q}_i^2}{2}$, $p_i = a_i^2 \dot{q}_i$, $H = \sum \frac{p_i^2}{2a_i^2} + U$. Donc,

$$H = p_\xi^2 \frac{(r_1 + r_2)^2 - 4c^2}{2r_1r_2} + p_\eta^2 \frac{4c^2 - (r_1 - r_2)^2}{2r_1r_2} - \frac{k}{r_1} - \frac{k}{r_2}.$$

Or, $r_1 + r_2 = \xi$, $r_1 - r_2 = \eta$, $4r_1r_2 = \xi^2 - \eta^2$ et donc finalement

$$H = 2p_\xi^2 \frac{\xi^2 - 4c^2}{\xi^2 - \eta^2} + 2p_\eta^2 \frac{4c^2 - \eta^2}{\xi^2 - \eta^2} - \frac{4k\xi}{\xi^2 - \eta^2}.$$

Résolvons maintenant l'équation de Hamilton-Jacobi.

Définition. Si dans l'équation

$$\Phi_1\left(\frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}; q_1, \ldots, q_n\right) = 0$$

la variable q_1 et la dérivée $\partial S/\partial q_1$ ne figurent que sous forme de combinaison $\varphi\left(\frac{\partial S}{\partial q_1}, q_1\right)$, on dit que la *variable q_1 est séparable*.

Dans ce cas il est utile de chercher les solutions de l'équation sous la forme

$$S = S_1(q_1) + S'(q_2, \ldots, q_n).$$

En posant $\varphi\left(\frac{\partial S_1}{\partial q_1}, q_1\right) = c_1$ dans l'équation initiale on obtient pour S' une équation avec un nombre moindre de variables

$$\Phi_2\left(\frac{\partial S'}{\partial q_2}, \ldots, \frac{\partial S'}{\partial q_n}, q_2, \ldots, q_n; c_1\right) = 0.$$

Soit $S' = S'(q_2, \ldots, q_n; c_1, c)$ une famille de solutions de cette équation dépendant des paramètres c_i. Les fonctions $S_1(q_1, c_1) + S'$ vérifieront l'équation initiale si la solution S_1 vérifie l'équation différentielle ordinaire $\varphi\left(\frac{dS_1}{dq_1}, q_1\right) = c_1$. Cette équation se résout

facilement; en effet, en exprimant $\frac{dS_1}{dq_1}$ en fonction de q_1 et c_1 on obtient $\frac{dS_1}{dq_1} = \psi(q_1, c_1)$, d'où $S_1 = \int^{q_1} \psi(q_1, c_1)\, dq_1$.

Si dans la nouvelle équation une variable est séparable, disons q_2, on peut répéter cette procédure et (si l'exemple le permet) trouver une solution de l'équation initiale dépendant de n constantes

$$S_1(q_1;\ c_1) + S_2(q_2;\ c_1,\ c_2) + \ldots + S_n(q_n;\ c_1, \ldots, c_n).$$

On dit dans ce cas que les variables sont *totalement séparables*.

Si les variables sont totalement séparables, alors des quadratures nous donneront une solution de l'équation de Hamilton-Jacobi $\Phi_1\left(\frac{\partial S}{\partial q}, q\right) = 0$ qui dépendra de n paramètres. Et l'on pourra intégrer par quadratures le système correspondant d'équations canoniques (théorème de Jacobi).

Appliquons ce qu'on a dit au problème de deux centres fixes. L'équation de Hamilton-Jacobi (4) s'écrit

$$\left(\frac{\partial S}{\partial \xi}\right)^2 (\xi^2 - 4c^2) + \left(\frac{\partial S}{\partial \eta}\right)^2 (4c^2 - \eta^2) = K(\xi^2 - \eta^2) + 4k\xi.$$

Nous pouvons séparer les variables en posant par exemple

$$\left(\frac{\partial S}{\partial \xi}\right)^2 (\xi^2 - 4c^2) - 4k\xi - K\xi^2 = c_1,$$

$$\left(\frac{\partial S}{\partial \eta}\right)^2 (4c^2 - \eta^2) + K\eta^2 = -c_1.$$

L'intégrale complète de l'équation (4) s'écrit alors

$$S(\xi, \eta;\ c_1, c_2) = \int \sqrt{\frac{c_1 + c_2\xi^2 + 4K\xi}{\xi^2 - 4c^2}}\, d\xi + \int \sqrt{\frac{-c_1 - c_2\eta^2}{4c^2 - \eta^2}}\, d\eta.$$

Le théorème de Jacobi nous donne maintenant explicitement l'expression du mouvement dans le problème de deux centres fixes en fonction d'intégrales elliptiques. Pour une étude plus détaillée de ce mouvement on pourra se référer à l'ouvrage de Charlier, *La mécanique céleste*, Paris, 1963.

On peut également appliquer le problème de l'attraction de deux centres fixes à l'étude des *mouvements sous une action constante dans le champ d'un centre attractif.*

Il s'agit ici du mouvement d'un point matériel sous l'action de l'attraction newtonienne d'un centre fixe et d'une forme (une « poussée ») constante en grandeur et en sens.

Ce problème peut être considéré comme un cas limite du problème de l'attraction de deux centres fixes. Lorsqu'on passe à la limite, le deuxième centre s'éloigne à l'infini dans le sens de la poussée (en outre sa masse doit grandir de manière à assurer une valeur

constante à la poussée, i.e. proportionnellement au carré de l'éloignement).

Ce cas limite du problème de l'attraction de deux centres fixes s'intègre explicitement (dans des fonctions elliptiques). On peut s'en assurer immédiatement par un passage à la limite ou par une séparation immédiate des variables dans le problème du mouvement sous une poussée constante dans le champ d'un seul centre.

Les coordonnées dans lesquelles les variables se séparent dans ce problème se déduisent des coordonnées elliptiques par passage à la limite lorsque l'un des centres s'éloigne à l'infini. On les appelle *coordonnées paraboliques* et on les définit par les formules

$$u = r - x, \quad v = r + x$$

(la poussée a même support et même sens que l'axe x).

On pourra trouver la description des trajectoires du mouvement sous une poussée constante (beaucoup d'entre elles sont très compliquées) dans l'ouvrage de V. Beletski, *Essais sur le mouvement des engins cosmiques* (en russe), Moscou, « Naouka », 1972.

Comme autre exemple étudions le problème des géodésiques sur un ellipsoïde à trois axes *). Utilisons les coordonnées elliptiques de Jacobi λ_1, λ_2, λ_3 qui sont racines de l'équation

$$\frac{x_1^2}{a_1+\lambda} + \frac{x_2^2}{a_2+\lambda} + \frac{x_3^2}{a_3+\lambda} = 1, \quad \lambda_1 > \lambda_2 > \lambda_3,$$

où x_1, x_2, x_3 sont des coordonnées cartésiennes. Nous ne ferons pas les calculs qui montrent que les variables sont séparables (on les trouvera par exemple dans les *Cours de dynamique* de Jacobi), nous nous limiterons simplement à la description des géodésiques.

Les surfaces $\lambda_1 = \text{const}$, $\lambda_2 = \text{const}$, $\lambda_3 = \text{const}$ sont des quadriques *homofocales*. L'une d'elles est un ellipsoïde, une autre un hyperboloïde à une nappe, la troisième un hyperboloïde à deux nappes. L'ellipsoïde peut dégénérer en l'intérieur d'une ellipse; l'hyperboloïde à une nappe, en l'extérieur d'une ellipse ou en la région du plan comprise entre les branches d'une hyperbole; l'hyperboloïde à deux nappes, en la région du plan située à l'extérieur des branches d'une hyperbole ou en un plan.

Supposons que l'ellipsoïde considéré fasse partie d'une famille d'ellipsoïdes de demi-axes $a > b > c$. Chacune des trois ellipses $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ est une géodésique fermée. Une géodésique issue des points d'une grande ellipse (de demi-axes a et b), dans une direction proche de celle de l'ellipse (fig. 206), est à tour de

---

*) Le problème des géodésiques sur l'ellipsoïde et le problème proche du billard ellipsoïdal ont trouvé une application dans de nombreuses recherches sur les appareils à laser.

rôle tangente aux deux lignes fermées d'intersection de l'ellipsoïde avec l'hyperboloïde à une nappe de notre famille λ = const *).

Une telle géodésique soit est fermée soit est partout dense dans la couronne comprise entre les lignes d'intersection. Plus la pente des géodésiques s'accroît, plus les hyperboloïdes se rapprochent du

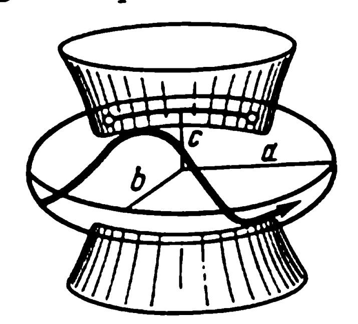


Fig. 206. Géodésique sur un ellipsoïde à trois axes.

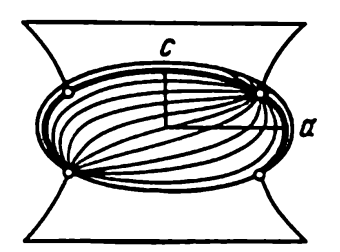


Fig. 207. Géodésiques issues d'un ombilic.

domaine « intérieur » à l'hyperbole qui coupe notre ellipsoïde en ses quatre « points ombilicaux ». Dans le cas limite on obtient des géodésiques passant par ces points (fig. 207).

Il est intéressant de remarquer que toutes les géodésiques issues d'un point ombilical se rassemblent dans le point ombilical opposé et qu'elles sont toutes de même longueur. Cependant une seule de

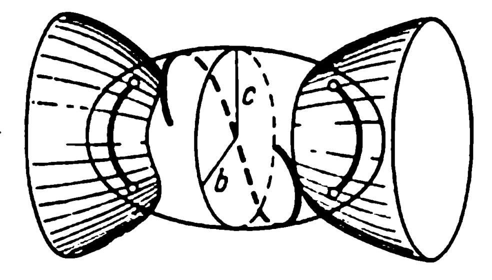


Fig. 208. Géodésiques d'un ellipsoïde tangentes à un hyperboloïde à deux nappes.

ces géodésiques est fermée : c'est l'ellipse moyenne de demi-axes a, c. Si l'on se déplace sur n'importe quelle autre géodésique passant par un point ombilical dans un sens ou dans l'autre, on se rapprochera asymptotiquement de cette ellipse moyenne.

Enfin, les géodésiques qui coupent la grande ellipse sous un angle proche de $\frac{\pi}{2}$ (fig. 208) sont tour à tour tangentes aux deux lignes d'intersection **) de notre ellipsoïde avec l'hyperboloïde à deux nappes. D'une façon générale, elles sont partout denses dans la couronne comprise entre ces lignes.

Parmi ces géodésiques il importe de distinguer la petite ellipse de demi-axes b, c.

*) Ces lignes d'intersection des quadriques homofocales sont également *lignes de courbure* de l'ellipsoïde.

**) Elles sont également lignes de courbure.

« La principale difficulté dans l'intégration de ces équations différentielles consiste à introduire des variables commodes. Or nous ne sommes guidés par aucune règle. Donc il nous faudra suivre le chemin inverse, c'est-à-dire trouver un changement remarquable et voir à quels problèmes il peut être appliqué avec succès » *) (J a c o b i, *Cours de dynamique*).

Dans la *Mécanique* de L. Landau et E. Lifchitz, Moscou, Ed. « Mir », 1964, au § 48 on trouvera une liste de problèmes admettant une séparation des variables en coordonnées sphériques, elliptiques et paraboliques.

§ 48. Fonctions génératrices

Dans ce paragraphe on bâtit l'appareil des fonctions génératrices pour les transformations canoniques non libres.

A. Fonction génératrice $S_2(P, q)$. Soit $g:\mathbf{R}^{2n}\to\mathbf{R}^{2n}$ une transformation canonique, $g(p, q) = (P, Q)$. Par définition de la transformation canonique, la 1-forme différentielle sur $\mathbf{R}^{2n}$

$$p\,dq - P\,dQ = dS$$

est différentielle totale d'une fonction $S(p, q)$. Une transformation canonique est *libre* si pour les $2n$ coordonnées on peut prendre q et Q. Dans ce cas S s'exprime en fonction de q et Q : on dit que c'est une *fonction génératrice* et on la note $S_1(q, Q)$. La connaissance de cette seule fonction permet de trouver les $2n$ fonctions qui définissent la transformation, à l'aide des relations

$$p = \frac{\partial S_1(q, Q)}{\partial q}, \quad P = -\frac{\partial S_1(q, Q)}{\partial Q}. \tag{1}$$

Cependant les transformations canoniques sont loin d'être toutes libres. Par exemple, dans le cas de la transformation identique, les coordonnées q et $Q = q$ sont dépendantes. Donc la transformation identique ne peut être définie par une fonction génératrice $S_1(q, Q)$.

Cependant on peut passer à une fonction génératrice d'une autre forme moyennant la transformation de Legendre. Supposons par exemple que pour coordonnées locales indépendantes sur $\mathbf{R}^{2n}$ on puisse prendre P, q $\left(\text{i. e. } \det\frac{\partial(P, q)}{\partial(p, q)} = \det\frac{\partial P}{\partial p}\right.$ est différent de $\left.\text{zéro}\right)$. Il vient

$$p\,dq - P\,dQ = dS, \quad p\,dq + Q\,dP = d(PQ + S).$$

*) Traduction libre du russe. — *Note du traducteur.*

La quantité $PQ + S$ exprimée en fonction de (P, q) est également une *fonction génératrice*:

$$S_2(P, q) = PQ + S(p, q).$$

On trouve pour cette fonction

$$p = \frac{\partial S_2(P, q)}{\partial q}, \quad Q = \frac{\partial S_2(P, q)}{\partial P}. \tag{2}$$

Inversement, si $S_2(P, q)$ est une fonction arbitraire telle que $\det \frac{\partial^2 S_2(P, q)}{\partial q\, \partial P}\Big|_{P_0, q_0} \neq 0$, alors au voisinage du point $\left(p_0 = \frac{\partial S_2(P, q)}{\partial q}\Big|_{P_0, q_0}, q_0\right)$ on peut résoudre le premier groupe d'équations (2) en P et obtenir les fonctions $P(p, q)$ (où $P(p_0, q_0) = P_0$). Puis le deuxième groupe d'équations (2) définit les fonctions $Q(p, q)$ et la transformation $(p, q) \to (P, Q)$ est canonique (prouvez-le!).

Exercice. Trouver la fonction génératrice S_2 de la transformation identique $P = p$, $Q = q$.

Réponse. Pq.

Remarque. La fonction génératrice $S_2(P, q)$ est commode parce que les formules (2) ne renferment pas de signes moins et on peut facilement les retenir, sachant que la fonction génératrice de la transformation identique est Pq.

B. 2^n-fonctions génératrices. Malheureusement les variables P et q ne peuvent pas toujours être prises pour coordonnées locales. Cependant on peut toujours faire choix d'une *certaine* collection de n nouvelles coordonnées

$$P_i = (P_{i_1}, \ldots, P_{i_k}), \quad Q_j = (Q_{j_1}, \ldots, Q_{j_{n-k}}),$$

de sorte qu'avec les anciens q on obtient $2n$ coordonnées indépendantes.

Ici $(i_1, \ldots, i_k)$ $(j_1, \ldots, j_{n-k})$ est une partition de l'ensemble $(1, \ldots, n)$ en deux sous-ensembles disjoints: ces partitions sont au nombre de 2^n.

Théorème. *Soit* $g: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ *une transformation canonique définie par les fonctions* $P(p, q)$, $Q(p, q)$. *Au voisinage de tout point* (p_0, q_0) *on peut prendre pour coordonnées indépendantes sur* $\mathbf{R}^{2n}$ *l'une au moins des* 2^n *collections de fonctions* (P_i, Q_j, q):

$$\det \frac{\partial (P_i, Q_j, q)}{\partial (p_i, p_j, q)} = \det \frac{\partial (P_i, Q_j)}{\partial (p_i, p_j)} \neq 0.$$

Au voisinage d'un tel point on peut retrouver la transformation canonique g *d'après la fonction*

$$S_3(P_i, Q_j, q) = (P_i, Q_i) + \int p\, dq - P\, dQ$$

moyennant les relations

$$p=\frac{\partial S_3}{\partial q}, \quad Q_i=\frac{\partial S_3}{\partial P_i}, \quad P_j=-\frac{\partial S_3}{\partial Q_j}. \tag{3}$$

Inversement, si $S_3(P_i, Q_j, q)$ *est une fonction quelconque telle que* $\det \frac{\partial^2 S_3}{\partial R \partial q}\Big|_{R_0, q_0}$ $(R = P_i, Q_j)$, *alors les expressions* (3) *définissent une transformation canonique au voisinage du point* p_0, q_0.

La démonstration de ce théorème est pratiquement la même que celle qui a été faite plus haut dans le cas particulier $k = n$. Il faut simplement vérifier que $\det \frac{\partial(P_i, Q_j)}{\partial(p_i, p_j)} \neq 0$ pour l'une des 2^n collections (P_i, Q_j, q).

Considérons la différentielle de la transformation g au point (p_0, q_0). En identifiant à $\mathbf{R}^{2n}$ les espaces qui lui sont tangents on peut estimer que dg est une transformation symplectique $S: \mathbf{R}^{2n} \to \mathbf{R}^{2n}$.

Considérons le p-plan P (i.e. un plan tel que $q = 0$) de $\mathbf{R}^{2n}$ (fig. 209). C'est un plan isotrope de dimension n et son image SP est également un plan isotrope.

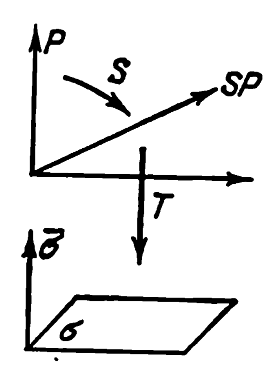


Fig. 209. Vérification de la non-dégénérescence.

Projetons le plan SP sur le plan arithmétique $\sigma = \{(p_i, q_j)\}$ parallèlement aux autres axes de coordonnées, i.e. suivant la direction du n-plan arithmétique isotrope $\bar\sigma = \{(p_j, q_i)\}$. Désignons par $T: SP \to \sigma$ l'opérateur de projection.

La condition $\det \frac{\partial(P_i, Q_j)}{\partial(p_i, p_j)} \neq 0$ signifie que $TS: P \to \sigma$ n'est pas dégénérée. L'opérateur S n'est pas dégénéré. Donc une condition nécessaire et suffisante pour que TS ne soit pas dégénérée est que la projection $T: SP \to \sigma$ ne le soit pas. En d'autres termes, le plan isotrope SP doit être transversal au plan arithmétique isotrope $\bar\sigma$. Or nous avons démontré au § 41 que l'un au moins des 2^n plans arithmétiques isotropes est transversal à SP. Donc l'un des 2^n déterminants est différent de zéro, c.q.f.d.

E x e r c i c e. Montrer que le système de 2^n types de fonctions génératrices que nous avons cité est minimal : il existe des transformations canoniques telles que soit différent de zéro l'un seulement des 2^n déterminants *).

C. Transformations canoniques infiniment petites. Considérons maintenant une transformation canonique voisine de la transfor-

*) Le nombre de types de fonctions génératrices proposées dans les divers ouvrages varie entre 4 et 4^n.

mation identique. Sa fonction génératrice peut être prise voisine de celle de l'identité Pq. Soit une famille de transformations canoniques g_ε dépendant différentiablement d'un paramètre ε. La fonction génératrice s'écrit

$$Pq+\varepsilon S(P, q; \varepsilon); \quad p=P+\varepsilon \frac{\partial S}{\partial q}, \quad Q=q+\varepsilon \frac{\partial S}{\partial P}. \tag{4}$$

On appelle *transformation canonique infiniment petite* la classe d'équivalence des familles g_ε ; deux familles g_ε et h_ε sont équivalentes si leur écart est un infiniment petit supérieur au premier ordre: $| g_\varepsilon - h_\varepsilon | = O(\varepsilon^2)$, $\varepsilon \to 0$.

Théorème. *Une transformation canonique infiniment petite vérifie les équations différentielles de Hamilton*

$$\left.\frac{dP}{d\varepsilon}\right|_{\varepsilon=0} = -\frac{\partial H}{\partial q}, \quad \left.\frac{dQ}{d\varepsilon}\right|_{\varepsilon=0} = \frac{\partial H}{\partial p}$$

de fonction de Hamilton $H(p, q) = S(p, q, 0)$.

La démonstration découle des formules (4): $P \to p$ pour $\varepsilon \to 0$.

Corollaire. *Une condition nécessaire et suffisante pour qu'un groupe à un paramètre de transformations de l'espace des phases* $\mathbf{R}^{2n}$ *vérifie les équations canoniques de Hamilton est que ces transformations soient canoniques.*

Pour cette raison la fonction de Hamilton H est appelée « fonction génératrice d'une transformation canonique infiniment petite ».

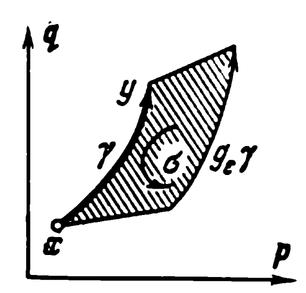


Fig. 210. Interprétation géométrique de la fonction de Hamilton.

Remarquons qu'à la différence des fonctions génératrices S, le hamiltonien H est une fonction de point de l'espace des phases intrinsèquement liée avec la transformation qu'elle définit.

La fonction H admet une interprétation géométrique simple. Soient x et y deux points de $\mathbf{R}^{2n}$ (fig. 210), γ une courbe passant par eux, $\partial\gamma = y - x$. Considérons les translations de la courbe γ par les transformations $g_\tau \gamma$, $0 \leqslant \tau \leqslant \varepsilon$. Elles forment une bande $\sigma(\varepsilon)$. Considérons l'intégrale de la forme $\omega^2 = \sum dp_i \wedge dq_i$ sur la 2-chaîne σ, $\partial\sigma = g_\varepsilon\gamma - \gamma + g_\tau y - g_\tau x$.

E x e r c i c e. Démontrer que

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\iint\limits_{\sigma(\varepsilon)}\omega^2 = H(y) - H(x)$$

existe et ne dépend pas du représentant de la classe g_ε.

De ce résultat on déduit une fois de plus le

Corollaire. *Les équations canoniques et la grandeur de la fonction de Hamilton sont invariantes par toute transformation canonique.*

En effet nous avons calculé l'accroissement de la fonction de Hamilton en utilisant seulement une transformation canonique infiniment petite et la structure symplectique de $\mathbf{R}^{2n}$, i.e. la forme ω^2.

CHAPITRE 10

INTRODUCTION À LA THÉORIE DES PERTURBATIONS

La théorie des perturbations est un arsenal très utile de méthodes de résolution approchée de problèmes « perturbés » voisins de problèmes « non perturbés » résolus exactement. Ces méthodes se justifient facilement quant à l'étude de mouvements sur un intervalle de temps court. Par contre on a très peu étudié la crédibilité de cette théorie dans le cas de mouvements s'effectuant sur des intervalles de temps grands ou infinis.

Nous verrons que dans de nombreux problèmes intégrables « non perturbés » le mouvement sera quasi périodique. Les variables « action-angle », qui sont des coordonnées symplectiques d'un type spécial, sont particulièrement commodes pour étudier les mouvements non perturbés et surtout les mouvements perturbés. En conclusion nous démontrerons le théorème fondamental de la théorie des perturbations des systèmes à fréquence unique et prouverons l'invariance adiabatique de la variable action dans de tels systèmes.

§ 49. Systèmes intégrables

Pour intégrer un système de $2n$ équations différentielles ordinaires il faut connaître $2n$ intégrales premières. Il s'avère que si est donné un système canonique d'équations différentielles, dans nombre de cas il suffit de connaître seulement n intégrales premières, chacune d'elles permettant d'abaisser l'ordre du système non pas de une mais de deux unités.

A. Théorème de Liouville sur les systèmes intégrables. Rappelons qu'une fonction F est intégrale première d'un système de hamiltonien H si et seulement si le crochet de Poisson est identiquement nul :

$$(H, F) \equiv 0.$$

D é f i n i t i o n. Deux fonctions F_1, F_2 définies sur une variété symplectique sont *en involution* si leur crochet de Poisson est nul.

Liouville a démontré que *si dans un système* à n *degrés de liberté* (*i. e. dont l'espace des phases est de dimension* $2n$) *on connaît* n *intégra-*

les premières indépendantes en involution, alors ce système s'intègre par quadratures.

L'énoncé exact de ce théorème est:

Soient données n fonctions en involution

$$F_1, \ldots, F_n; \quad (F_i, F_j) \equiv 0, \; i, j = 1, 2, \ldots, n,$$

sur une variété symplectique de dimension 2n. Considérons l'ensemble de niveau des fonctions F_i

$$M_f = \{x : F_i(x) = f_i, \; i = 1, \ldots, n\}.$$

Supposons que sur M_f *les n fonctions* F_i *sont indépendantes (i.e. les n 1-formes* dF_i *sont linéairement indépendantes en chaque point de* M_f*).*

Alors:

1) M_f *est une variété différentiable invariante par le flot de hamiltonien* $H = F_1$.

2) *Si la variété* M_f *est compacte et connexe, alors elle est difféomorphe à un tore de dimension n*

$$T^n = \{(\varphi_1, \ldots, \varphi_n) \bmod 2\pi\}.$$

3) *Le flot de hamiltonien H définit sur* M_f *un mouvement quasi périodique, i.e. en coordonnées angulaires* $\varphi = (\varphi_1, \ldots, \varphi_n)$

$$\frac{d\varphi}{dt} = \omega, \quad \omega = \omega(f).$$

4) *Les équations canoniques de hamiltonien H sont intégrables par quadratures.*

Avant de démontrer ce théorème énonçons quelques-uns de ses corollaires.

Corollaire 1. *Si dans un système canonique à deux degrés de liberté on connaît une intégrale première F ne dépendant pas du hamiltonien H, alors ce système s'intègre par quadratures; la sous-variété compacte connexe* $H = h$, $F = f$ *de dimension deux de l'espace des phases est un tore invariant sur lequel le mouvement est quasi périodique.*

En effet, F et H sont en involution, puisque F est intégrale première du système de hamiltonien H.

Comme exemple à trois degrés de liberté considérons une toupie lagrangienne symétrique pesante fixée en un point de son axe. On déduit immédiatement trois intégrales premières: H, M_z, M_3. Il est aisé de vérifier que les intégrales M_z et M_3 sont en involution. D'autre part, la variété $H = h$ est compacte dans l'espace des phases. Donc sans effectuer le moindre calcul on peut dire que pour la plupart des conditions initiales *) le mouvement de la toupie est

*) A l'exception d'ensembles singuliers de niveau d'intégrales où l'indépendance n'a pas lieu.

quasi périodique: les trajectoires de phase sont partout denses sur les tores de dimension trois $H = c_1$, $M_z = c_2$, $M_3 = c_3$. Les trois fréquences correspondantes sont appelées fréquences de rotation propre, de précession et de nutation.

Les autres exemples résultent de la remarque suivante: *si un système canonique est intégrable par la méthode de Jacobi-Hamilton, il possède n intégrales premières en involution.*

En effet, la méthode consiste à trouver une transformation canonique $(p, q) \to (P, Q)$ telle que P_i soient intégrales premières. Or les fonctions P_i et P_j sont visiblement en involution.

En particulier, ce qui précède est valable pour le problème de l'attraction par deux centres fixes. Il est aisé de multiplier les exemples; le théorème de Liouville énoncé plus haut englobe pratiquement tous les problèmes de dynamique intégrés à ce jour.

B. Début de la démonstration du théorème de Liouville. Passons maintenant à la démonstration du théorème. Considérons l'ensemble de niveau d'intégrales

$$M_f = \{x : F_i = f_i,\ i = 1, \ldots, n\}.$$

Par hypothèse, les n 1-formes dF_i sont linéairement indépendantes en chaque point de M_f. Donc, d'après le théorème des fonctions implicites, M_f est une sous-variété de dimension n de l'espace des phases de dimension $2n$.

Lemme 1. *Sur la variété M_f de dimension n existent n champs de vecteurs tangents commutant deux à deux et linéairement indépendants en chaque point.*

D é m o n s t r a t i o n. La structure symplectique de l'espace des phases définit un opérateur I qui transforme les 1-formes en champs de vecteurs. Cet opérateur transforme la 1-forme dF_i en le champ $I\,dF_i$ de la vitesse de phase du système de hamiltonien F_i. Montrons que les *n champs $I\,dF_i$ sont tangents à M_f, commutent et sont indépendants.*

En effet, l'indépendance des dF_i et la non-dégénérescence de l'isomorphisme I entraînent l'indépendance des $I\,dF_i$ en chaque point de M_f. Les champs $I\,dF_i$ commutent deux à deux, puisque les crochets de Poisson de leurs hamiltoniens $(F_i, F_j) \equiv 0$. Pour la même raison la dérivée de la fonction F_i suivant la direction du champ $I\,dF_j$ est nulle quels que soient $i, j = 1, \ldots, n$. Donc les champs $I\,dF_i$ sont tangents à M_f et le lemme 1 est démontré.

On remarquera qu'on a démontré plus que le lemme 1:

1′) *La variété M_f est invariante par chacun des n flots g_i^t, commutant deux à deux, de hamiltoniens F_i: $g_i^t g_j^s = g_j^s g_i^t$.*

1″) *La variété M_f est nulle (i.e. la 2-forme ω^2 est nulle sur $TM_f|_x$).*

En effet, les n vecteurs $I\, dF_i\, |_x$ sont deux à deux orthogonaux gauches ($(F_i, F_j) \equiv 0$) et forment une base dans le plan tangent à la variété M_f en x.

C. Variétés sur lesquelles opère transitivement un groupe de $\mathbf{R}^n$. Utilisons maintenant la proposition de topologie suivante.

Lemme 2. *Soit M^n une variété différentiable compacte connexe de dimension n sur laquelle sont définis n champs de vecteurs commutant deux à deux et linéairement indépendants en chaque point de M^n. Alors la variété M^n est difféomorphe à un tore de dimension n.*

Démonstration*). Désignons par g_i^t, $i = 1, \ldots, n$, les groupes à un paramètre de difféomorphismes de M correspondant aux n champs de vecteurs donnés. Comme ces champs commutent, les groupes g_i^t, g_j^s commutent également. Nous pouvons donc définir l'action g d'un groupe commutatif de $\mathbf{R}^n = \{t\}$ sur la variété M en posant

$$g^t : M \to M, \quad g^t = g_1^{t_1} \ldots g_n^{t_n} \quad (t = (t_1, \ldots, t_n) \in \mathbf{R}^n).$$

De toute évidence $g^{t+s} = g^t g^s \; \forall\, t, s \in \mathbf{R}^n$. Fixons un point $x_0 \in M$. Définissons l'application

$$g : \mathbf{R}^n \to M, \qquad g(t) = g^t x_0$$

(le point x_0 doit être déplacé de t_1 sur une trajectoire du premier flot, de t_2 sur une trajectoire du second, etc.).

Exercice 1. Montrer que l'*application g* (fig. 211) *a un voisinage suffisamment petit V du point $O \in \mathbf{R}^n$ définit une carte du voisinage du point x_0: chaque point $x_0 \in M$ possède un voisinage U, $x_0 \in U \subset M$ tel que g soit un difféomorphisme de V sur U.*

Indication. Appliquer le théorème des fonctions implicites et se servir de l'indépendance linéaire des champs au point x_0.

Exercice 2. Montrer que $g : \mathbf{R}^n \to M$ est une *application sur.*

Indication. Relier un point $x \in M$ à x_0 par une courbe (fig. 212), recouvrir cette courbe par un nombre fini de voisinages U de l'exercice précédent et définir t comme une somme de t_i-translations correspondant aux morceaux de courbe.

Remarquons que l'application $g : \mathbf{R}^n \to M^n$ ne peut être bijective, puisque M^n est compact et $\mathbf{R}^n$ non. Etudions l'ensemble des contre-images du point $x_0 \in M^n$.

Définition. On appelle *groupe stationnaire* d'un point x_0 l'ensemble G des points $t \in \mathbf{R}^n$ tels que $g^t x_0 = x_0$.

Exercice 3. Montrer que *G est un sous-groupe de $\mathbf{R}^n$ ne dépendant pas du point x_0.*

---

*) La démonstration sera achevée à la page 276.

S o l u t i o n. Si $g^t x_0 = x_0$, $g^s x_0 = x_0$, alors

$$g^{t+s}x_0 = g^s g^t x_0 = g^s x_0 = x_0, \quad g^{-t}x_0 = g^{-t} g^t x_0 = x_0.$$

Donc le groupe G est un sous-groupe de $\mathbf{R}^n$. Si $x = g^\tau x_0$, $t \in G$, alors $g^t x = g^{t+\tau} x = g^\tau g^t x_0 = g^\tau x_0 = x$.

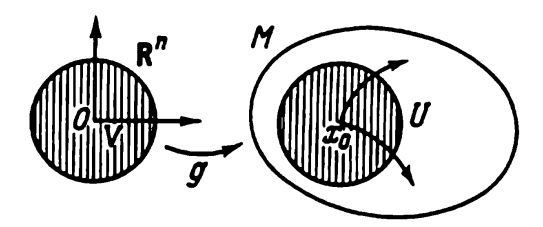


Fig. 211. Problème 1.

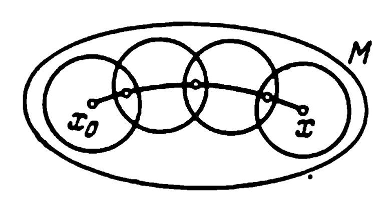


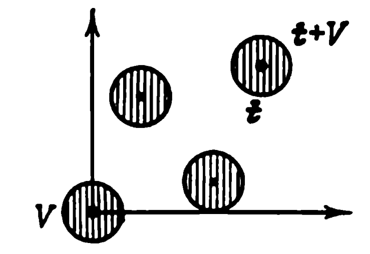
Fig. 212. Problème 2.

Donc le groupe stationnaire G est un sous-groupe de $\mathbf{R}^n$ entièrement défini et ne dépendant pas de x_0. En particulier, le point $t = 0$ appartient de toute évidence à G.

E x e r c i c e 4. Montrer que *dans un voisinage V suffisamment petit du point* $t = 0$ de $\mathbf{R}^n$ *le sous-groupe stationnaire G ne possède pas d'autre point en dehors de* $t = 0$.



Fig. 213. Problème 5.



I n d i c a t i o n. L'application $g: V \to U$ est un difféomorphisme.

E x e r c i c e 5. Montrer qu'un voisinage $t + V$ (fig. 213) *d'un point quelconque* $t \in G \subset \mathbf{R}^n$ *et le groupe stationnaire G n'ont qu'un point commun*: *le point* t.

Donc les points du sous-groupe stationnaire G ne possèdent pas de points d'accumulation dans $\mathbf{R}^n$. De tels sous-groupes sont appelés *sous-groupes discrets*.

E x e m p l e. Soient $e_1, \ldots, e_k$ k vecteurs linéairement indépendants de $\mathbf{R}^n$, $0 \leqslant k \leqslant n$. L'ensemble de toutes les combinaisons linéaires entières (fig. 214)

$$m_1 e_1 + \ldots + m_k e_k, \quad m_i \in \mathbf{Z} = (\ldots -2, -1, 0, 1, \ldots),$$

forme un sous-groupe discret de $\mathbf{R}^n$. L'ensemble de tous les points entiers du plan est un sous-groupe discret du plan.

D. Sous-groupes discrets de $\mathbf{R}^n$. L'exemple cité englobe tous les sous-groupes discrets de $\mathbf{R}^n$. Plus exactement on démontrera le

Lemme 3. *Soit G un sous-groupe discret de $\mathbf{R}^n$. Il existe alors k vecteurs $e_1, \ldots, e_k \in G$ $(0 \leqslant k \leqslant n)$ linéairement indépendants tels que G soit l'ensemble de toutes leurs combinaisons linéaires entières.*

D é m o n s t r a t i o n. Envisageons une structure euclidienne sur $\mathbf{R}^n$. On a toujours $O \in G$. Si $G = \{O\}$, le lemme est démontré. Sinon il existe un point $e_0 \in G$, $e_0 \neq O$ (fig. 215). Considérons la

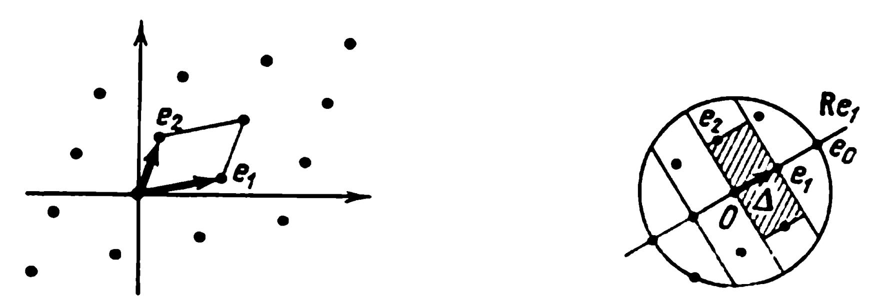


Fig. 214. Sous-groupe discret du plan. Fig. 215. Démonstration du lemme 3.

droite $\mathbf{R}e_0$. Montrons que cette droite contient le point e_1 le plus proche de O. En effet, la boule de rayon $| e_0 |$ et de centre O ne contient qu'un nombre fini de points de G (chaque point de G, comme nous l'avons vu plus haut, possède un voisinage tel que V ne comprenant pas d'autres points de G).

Du nombre fini de points de cette boule, situés sur la droite $\mathbf{R}e_0$, le plus proche de O sera précisément le point de G le plus proche de O.

Les seuls points de $\mathbf{R}e_0$ à appartenir au groupe G sont les multiples entiers de e_1 $(me_1, m \in \mathbf{Z})$. En effet, les points me_1 divisent la droite en parties de longueur $| e_1 |$. Si l'une de ces parties $(me_1, (m+1)\, e_1)$ contenait le point $e \in G$, alors le point $e - me_1 \in G$ serait plus proche de O que e_1.

Si le groupe G ne possède pas de points en dehors de la droite $\mathbf{R}e_1$, le lemme est démontré. Supposons qu'il existe un point $e \in G$, $e \notin \mathbf{R}e_1$. Montrons qu'il existe un point $e_2 \in G$ qui est le plus proche de $\mathbf{R}e_1$ (ce point n'est pas situé sur la droite). Projetons e orthogonalement sur la droite $\mathbf{R}e_1$. La projection couvre entièrement un segment $\Delta = \{\lambda e_1\}$, $m \leqslant \lambda < m + 1$. Considérons le cylindre circulaire droit d'axe Δ et de rayon égal à la distance de Δ à e. Ce cylindre renferme un nombre fini (non nul) de points du groupe G.

Soit e_2 le point le plus proche de l'axe $\mathbf{R}e_1$ non situé sur cet axe.

E x e r c i c e 6. Montrer que la *distance d'un point quelconque du groupe G non situé sur l'axe $\mathbf{R}e_1$ à cet axe n'est pas inférieure à la distance du point e_2 à l'axe $\mathbf{R}e_1$.*

I n d i c a t i o n. Moyennant une me_1-translation on peut faire en sorte que la projection sur l'axe tombe sur le segment Δ.

Les combinaisons linéaires entières de e_1 et e_2 forment un treillis dans le plan $\mathbf{R}e_1 + \mathbf{R}e_2$.

E x e r c i c e 7. Montrer que le *plan* $Re_1 + Re_2$ *ne contient pas de points du groupe G autres que les combinaisons linéaires de* e_1 *et* e_2.

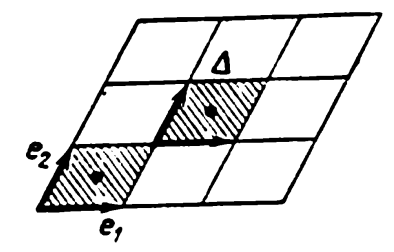


Fig. 216. Problème 7.

I n d i c a t i o n. Partager le plan en parallélogrammes (fig. 216) $\Delta = \{\lambda_1 e_1 + \lambda_2 e_2\}$, $m_i \leqslant \lambda_i < m_i + 1$. Si $e \in \Delta$, $e \neq m_1 e_1 + m_2 e_2$, alors le point $e - m_1 e_1 - m_2 e_2$ serait plus proche de Re_1 que e_2.

Si le groupe G ne possède pas de points non situés sur le plan $\mathbf{R}e_1 + \mathbf{R}e_2$, le lemme est démontré. Supposons qu'il existe un point $e \in G$ non contenu dans ce plan. Alors il existe un point $e_3 \in G$ qui est le plus proche du plan $\mathbf{R}e_1 + \mathbf{R}e_2$; les points $m_1 e_1 + m_2 e_2 + m_3 e_3$ représentent l'intersection de G avec l'espace $\mathbf{R}e_1 + \mathbf{R}e_2 + \mathbf{R}e_3$; s'il n'en est pas ainsi, on prend le plus proche point de cet espace, etc.

E x e r c i c e 8. Montrer que *le plus proche point existe toujours.*
I n d i c a t i o n. Prendre le plus proche du nombre fini de points du « cylindre » correspondant.

Remarquons que tous les vecteurs $e_1, e_2, e_3, \ldots$ obtenus sont linéairement indépendants. Comme ils appartiennent tous à $\mathbf{R}^n$, leur nombre k est inférieur à n.

E x e r c i c e 9. Montrer que *G est composé exactement de toutes les combinaisons linéaires entières de* $e_1, \ldots, e_k$.
I n d i c a t i o n. Partager le plan $Re_1 + \ldots + Re_k$ en parallélépipèdes Δ et montrer qu'aucun Δ ne contient des points de G. S'il existe un point $e_1 \in G$ non contenu dans le plan $Re_1 + \ldots + Re_k$, la construction n'est pas achevée.

Donc le lemme 3 est démontré.

Maintenant il n'est pas difficile de démontrer le lemme 2: M_f est difféomorphe à un tore T^n.

Considérons le produit cartésien de k cercles et $n - k$ droites:

$$T^k \times \mathbf{R}^{n-k} = \{(\varphi_1, \ldots, \varphi_k; y_1, \ldots, y_{n-k})\}, \ \varphi \bmod 2\pi,$$

et l'application naturelle $p: \mathbf{R}^n \to T^k \times \mathbf{R}^{n-k}$

$$p(\varphi, y) = (\varphi \bmod 2\pi, y).$$

Les points $f_1, \ldots, f_k \in \mathbf{R}^n$ (f_i a pour coordonnées $\varphi_i = 2\pi$, $\varphi_j = 0$, $y = 0$) sont envoyés en 0 par cette application.

Supposons que $e_1, \ldots, e_k \in G \subset \mathbf{R}^n$ sont générateurs du groupe stationnaire G (voir lemme 3). Appliquons l'espace vectoriel $\mathbf{R}^n =$

$= \{(\varphi, y)\}$ sur l'espace $\mathbf{R}^n = \{t\}$, de telle sorte que les vecteurs f_i se transforment en e_i. Soit $A: \mathbf{R}^n \to \mathbf{R}^n$ un tel isomorphisme.

Remarquons maintenant que $\mathbf{R}^n = \{(\varphi, y)\}$ définit les cartes de $T^k \times \mathbf{R}^{n-k}$ et $\mathbf{R}^n = \{t\}$, les cartes de la variété M_f.

Exercice 10. Montrer que l'*application des cartes* $A: \mathbf{R}^n \to \mathbf{R}^n$ *définit un difféomorphisme* $\widetilde{A}: T^k \times \mathbf{R}^{n-k} \to M_f$,

$$\begin{array}{ccc} \mathbf{R}^n = \{(\varphi, y)\} & \xrightarrow{A} & \mathbf{R}^n = \{t\} \\ p \downarrow & & \downarrow g \\ T^k \times \mathbf{R}^{n-k} & \xrightarrow{\widetilde{A}} & M_f \end{array} .$$

Or comme par hypothèse la variété M_f est compacte, $k = n$ et M_f est un tore de dimension n. Le lemme 2 est démontré.

En vertu du lemme 1 sont donc démontrées les deux premières propositions du théorème de Liouville.

Par la même occasion nous avons construit sur M_f les coordonnées angulaires $\varphi_1, \ldots, \varphi_n$ mod 2π.

Exercice 11. *Montrer que sous l'action du flot de hamiltonien H les coordonnées angulaires φ varient uniformément*:

$$\dot{\varphi}_i = \omega_i; \quad \omega_i = \omega_i(f); \quad \varphi(t) = \varphi(0) + \omega t.$$

En d'autres termes, le *mouvement est quasi périodique sur le tore invariant* M_f.
Indication. $\varphi = A^{-1}t$.

De toutes les propositions du théorème il reste à démontrer la dernière, à savoir que le système s'intègre par quadratures.

§ 50. Les variables action-angle

On montrera ici que les conditions du théorème de Liouville étant réalisées, il est possible de choisir des coordonnées symplectiques (I, φ) telles que les intégrales premières F dépendent uniquement de I, φ étant des coordonnées angulaires sur le tore M_f.

A. Description des variables action-angle. Au § 49 nous avons étudié exclusivement la variété compacte connexe de niveau d'intégrales: $M_f = \{x: F(x) = f\}$; il s'est avéré que M_f est un tore de dimension n invariant par le flot. Nous avons choisi des coordonnées angulaires φ_i sur M_f telles que le flot de hamiltonien $H = F_1$ prenne sur M_f la forme particulièrement simple

$$\frac{d\varphi}{dt} = \omega(f), \quad \varphi(t) = \varphi(0) + \omega t.$$

Considérons maintenant un voisinage de la variété M_f dans l'espace des phases de dimension $2n$.

Exercice. Démontrer que la *variété M_f possède un voisinage difféomorphe au produit cartésien d'un tore T^n par une boule D^n d'un espace euclidien de dimension n.*

Indication. Prendre pour coordonnées les fonctions F_i et les angles φ_i construits ci-dessus. Les dF_i étant linéairement indépendants, les fonctions F_i et φ_i $(i = 1, \ldots, n)$ définissent un difféomorphisme d'un voisinage de M_f sur le produit cartésien $T^n \times D^n$.

Dans les coordonnées (F, φ) introduites, le flot de hamiltonien $H = F_1$ s'écrit sous forme particulièrement simple d'un système de $2n$ équations différentielles ordinaires

$$\frac{dF}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega(F), \tag{1}$$

qui s'intègre immédiatement: $F(t) = F(0)$, $\varphi(t) = \varphi(0) + \omega(F(0))\,t$.

Donc pour intégrer explicitement le système canonique initial d'équations différentielles il suffit de trouver explicitement les variables φ. On peut le faire moyennant des quadratures. Plus bas on indique comment construire les variables φ.

Remarquons que les variables (F, φ) ne sont pas d'une façon générale des coordonnées symplectiques. Il s'avère qu'il existe des fonctions de F, désignons-les par $I = I(F)$, $I = (I_1, \ldots, I_n)$ telles que les variables (I, φ) soient a priori des coordonnées symplectiques: la structure symplectique initiale ω^2 s'exprime en fonction de ces coordonnées d'après la formule connue

$$\omega^2 = \sum dI_i \wedge d\varphi_i.$$

Les variables I sont appelées *variables d'action* *); avec les variables angulaires φ elles forment dans un voisinage de la variété M_f un système de coordonnées canoniques *action-angle*.

Les grandeurs I_i sont intégrales premières du système de hamiltonien $H = F_1$ puisqu'elles sont fonctions des intégrales premières F_j. A leur tour les variables F_i peuvent être exprimées en fonction de I et en particulier $H = F_1 = H(I)$. Dans les variables action-angle les équations différentielles du flot (1) prennent la forme

$$\frac{dI}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega(I). \tag{2}$$

Exercice. La fonction $\omega(I)$ de (2) peut-elle être arbitraire?

Solution. Dans les variables (I, φ) les équations du flot (2) s'écrivent sous la forme canonique de hamiltonien $H(I)$. Donc $\omega(I) = \partial H/\partial I$; si le nombre de degrés de liberté $n \geqslant 2$, les fonctions $\omega(I)$ ne sont pas arbitraires, elles vérifient la condition de symétrie $\partial\omega_i/\partial I_j = \partial\omega_j/\partial I_i$.

Les variables action-angle sont particulièrement importantes pour la théorie des perturbations; dans le § 52 on indique comment on les applique à la théorie des invariants adiabatiques.

*) Il est clair que I est de même dimension que l'action.

B. Construction des variables action-angle dans le cas d'un seul degré de liberté. Un système à un degré de liberté est donné sur le plan des phases $\{(p, q)\}$ par le hamiltonien $H(p, q)$.

Exemple 1. L'oscillateur harmonique $H = 1/2\ p^2 + 1/2\ q^2$ ou en général $H = 1/2\ a^2p^2 + 1/2\ b^2q^2$.

Exemple 2. Le pendule mathématique $H = 1/2p^2 - \cos q$.

Dans les deux cas nous avons des courbes fermées compactes M_h $(H = h)$ et nous sommes placés dans les conditions du théorème du § 49 pour $n = 1$.

Pour construire les variables action-angle nous chercherons une transformation canonique $(p, q) \to (I, \varphi)$ vérifiant les deux conditions suivantes:

$$1)\ I = I(h), \quad 2) \oint_{M_h} d\varphi = 2\pi. \tag{3}$$

Exercice. Trouver les variables action-angle dans le cas d'un simple oscillateur harmonique $H = 1/2\ p^2 + 1/2\ q^2$.

Solution. Si r et φ sont les coordonnées polaires, alors $dp \wedge dq = r\,dr \wedge d\varphi = d\frac{r^2}{2} \wedge d\varphi$. Donc $I = H = \frac{p^2+q^2}{2}$.

Pour construire une transformation canonique $p, q \to I, \varphi$ dans le cas général nous chercherons sa fonction génératrice $S(I, q)$:

$$p = \frac{\partial S(I, q)}{\partial q}, \quad \varphi = \frac{\partial S(I, q)}{\partial I}; \quad H\left(\frac{\partial S(I, q)}{\partial q}, q\right) = h(I). \tag{4}$$

Supposons tout d'abord que la fonction $h(I)$ est connue et inversible, de sorte que toute courbe M_h est définie par une valeur de I $(M_h = M_{h(I)})$. Alors pour une valeur fixe de I il vient à partir de (4)

$$dS\,|_{I=\text{const}} = p\,dq.$$

Cette relation définit sur la courbe $M_{h(I)}$ la 1-forme différentielle dS.

En intégrant cette 1-forme le long de la courbe $M_{h(I)}$ on obtient au voisinage d'un point q_0 la fonction

$$S(I, q) = \int_{q_0}^{q} p\,dp.$$

Cette fonction sera précisément la fonction génératrice de la transformation (4) au voisinage du point (I, q_0). La première des conditions (3) est automatiquement réalisée: $I = I(h)$. Pour satisfaire la deuxième condition voyons comment $S(I, q)$ se comporte globalement.

En parcourant une fois la courbe fermée $M_{h(I)}$, on donne à l'intégrale de $p\,dq$ l'accroissement

$$\Delta S(I) = \oint_{M_{h(I)}} p\,dq,$$

égal à l'aire A interceptée par la courbe $M_{h(I)}$. Donc la fonction S est une « fonction multivoque » sur la courbe $M_{h(I)}$: elle est définie à l'addition près d'un multiple de A. Ce terme n'a aucun effet sur $\frac{\partial S(I, q)}{\partial q}$; cependant il entraîne la multivocité de $\varphi = \partial S/\partial I$. Cette dérivée s'avère définie seulement à l'addition près d'un multiple de $\frac{d}{dI}\Delta S(I)$. D'une façon plus précise, les formules (4) définissent la 1-forme $d\varphi$ sur la courbe $M_{h(I)}$. L'intégrale de cette forme le long de $M_{h(I)}$ est égale à $\frac{d}{dI}\Delta S(I)$.

Pour que la deuxième de nos conditions soit réalisée, i.e. $\oint_{M_h} d\varphi = 2\pi$, il est nécessaire que

$$\frac{d}{dI}\Delta S(I) = 2\pi, \quad I = \frac{\Delta S}{2\pi} = \frac{A}{2\pi},$$

où $A = \oint_{M_h} p\,dq$ est l'aire interceptée par la courbe de phase $H = h$.

Définition. On appelle *variable d'action* dans le problème unidimensionnel de hamiltonien $H(p, q)$ la quantité $I(h) = 1/2\pi\, A(h)$.

Nous sommes en fin de compte arrivés à la conclusion suivante. Soit $dA/dh \neq 0$. Alors est définie la fonction $h(I)$ inverse de $I(h)$.

Théorème. *Posons* $S(I, q) = \int_{q_0}^{q} p\,dq\,|_{H=h(I)}$. *Alors les formules* (4) *définissent une transformation canonique* $p, q \to I, \varphi$ *vérifiant les conditions* (3).

Ceci achève la construction des variables action-angle dans le cas unidimensionnel.

Exercice. Trouver S et I pour l'oscillateur harmonique.

Réponse. Si $H = \frac{1}{2}a^2p^2 + \frac{1}{2}b^2q^2$ (fig. 217), alors M_h est une ellipse interceptant l'aire $A(h) = \pi\frac{\sqrt{2h}}{a}\frac{\sqrt{2h}}{b} = \frac{2\pi h}{ab} = \frac{2\pi h}{\omega}$. Donc, *s'agissant de l'oscillateur harmonique, la variable d'action est le quotient de l'énergie par la fréquence.* La variable angle est de toute évidence la phase d'oscillations.

E x e r c i c e. Montrer que la *période T du mouvement sur la courbe fermée $H = h$ du plan des phases $\{(p, q)\}$ est égale à la dérivée par rapport à h de l'aire interceptée par cette courbe*

$$T = \frac{dA(h)}{dh}.$$

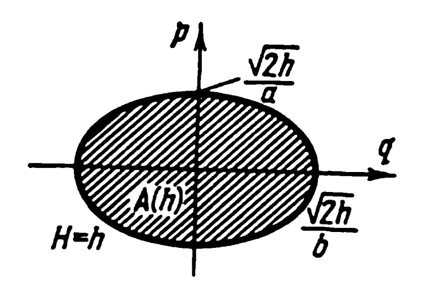


Fig. 217. Variable d'action pour un oscillateur harmonique.

S o l u t i o n. Dans les variables action-angle les équations du mouvement (2) s'écrivent

$$\dot{\varphi} = \frac{\partial H}{\partial I} = \left(\frac{dI}{dh}\right)^{-1} = 2\pi \left(\frac{dA}{dh}\right)^{-1}, \quad T = \frac{2\pi}{\dot{\varphi}} = \frac{dA}{dh}.$$

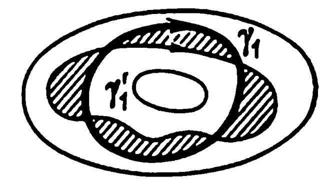
C. Construction des variables action-angle dans $\mathbf{R}^{2n}$. Passons maintenant à un système à n degrés de liberté défini sur $\mathbf{R}^{2n} = \{(p, q)\}$ par le hamiltonien $H(p, q)$ et possédant n intégrales premières en involution $F_1 = H, F_2, \ldots, F_n$. Sans répéter les raisonnements qui nous ont conduits au choix de $2\pi I = \oint p\, dq$ dans le cas unidimensionnel, définissons immédiatement n variables action I.

Soient $\gamma_1, \ldots, \gamma_n$ des cycles de base unidimensionnels du tore M_f (l'accroissement d'une coordonnée φ_i sur un cycle γ_i est égal à 2π si $i = j$ et à 0 si $i \neq j$). Posons

$$I_i(f) = \frac{1}{2\pi} \oint_{\gamma_i} p\, dq. \tag{5}$$

E x e r c i c e. *Montrer que cette intégrale ne dépend pas du choix de la courbe γ_i du cycle de base* (fig. 218).

Fig. 218. Indépendance de la variable d'action par rapport au choix de la courbe d'intégration.



I n d i c a t i o n. Au § 49 on a montré que la 2-forme $\omega^2 = \sum dp_i \wedge dq_i$ était nulle sur la variété M_f. La formule de Stokes donne

$$\oint_{\gamma} - \oint_{\gamma'} p\, dq = \iint_{\sigma} dp \wedge dq = 0,$$

où $\partial\sigma = \gamma - \gamma'$.

Définition. On appelle *variable d'action* les n grandeurs $I_i(f)$ définies par les formules (5).

Supposons maintenant que les n grandeurs I_i sont indépendantes pour des valeurs f_i données des n intégrales F_i: $\det \frac{\partial I}{\partial f}\Big|_f \neq 0$. Alors au voisinage du tore M_f on peut prendre pour coordonnées les variables I, φ.

Théorème. *La transformation* $p, q \to I, \varphi$ *est canonique, i.e.*

$$\sum dp_i \wedge dq_i = \sum dI_i \wedge d\varphi_i.$$

Traçons la ligne générale de cette démonstration. Soit la 1-forme différentielle $p\,dq$ sur M_f. La variété M_f étant lagrangienne (§ 49),

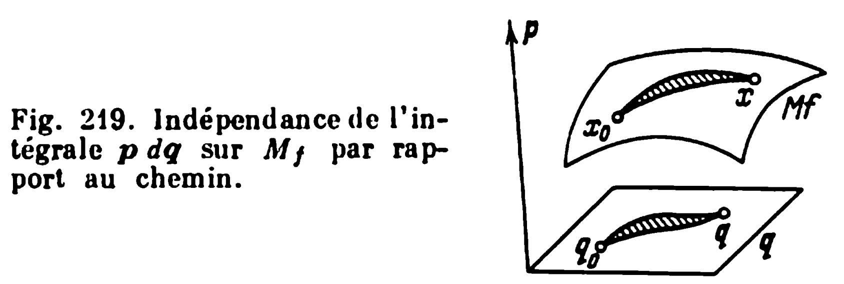


Fig. 219. Indépendance de l'intégrale $p\,dq$ sur M_f par rapport au chemin.

cette 1-forme est fermée dans M_f: sa dérivée extérieure $\omega^2 = dp \wedge dq$ sur M_f est identiquement nulle. Donc (fig. 219)

$$S(x) = \int_{x_0}^{x} p\,dq\,\Big|_{M_f}$$

ne change pas si le chemin d'intégration est déformé (formule de Stokes). Donc $S(x)$ est une « fonction multivoque » sur M_f; ses périodes sont égales à

$$\Delta_i S = \oint_{\gamma_i} dS = 2\pi I_i.$$

Soit maintenant x_0 un point de M_f au voisinage duquel n variables q font office de coordonnées sur M_f, de sorte que la sous-variété $M_f \subset \mathbf{R}^{2n}$ est définie par n équations de la forme $p = p(I, q)$, $q(x_0) = q_0$. Dans un voisinage simplement connexe du point q_0 est définie la fonction univoque

$$S(I, q) = \int_{q_0}^{q} p(I, q)\,dq,$$

que nous pouvons prendre pour fonction génératrice de la transformation canonique $p, q \to I, \varphi$:

$$p = \frac{\partial S}{\partial q}, \quad \varphi = \frac{\partial S}{\partial I}.$$

Il est aisé de vérifier que ces formules définissent réellement une transformation canonique non seulement au voisinage du point envisagé mais globalement au voisinage de M_f et que les coordonnées φ seront des fonctions multivoques de périodes

$$\Delta_i\varphi_j = \Delta_i \frac{\partial S}{\partial I_j} = \frac{\partial}{\partial I_j} \Delta_i S = \frac{\partial}{\partial I_j} 2\pi I_i = 2\pi\delta_{ij},$$

c.q.f.d.

On remarquera que toutes nos constructions ne contiennent que des opérations « algébriques » (inversion de fonctions) et une « quadrature ». Donc un système canonique de $2n$ équations dont on connaît n intégrales en involution se résout par quadratures, ce qui démontre la dernière affirmation du théorème de Liouville (§ 49).

R e m a r q u e 1. Dans le cas unidimensionnel déjà les variables action-angle I, φ ne sont pas définies univoquement par les conditions (3). Plus exactement, pour variable d'action on aurait pu prendre $I' = I +$ const et pour variable angulaire $\varphi' = \varphi + c\,(I)$.

R e m a r q u e 2. Nous avons construit les variables action-angle pour un système dont l'espace des phases est $\mathbf{R}^{2n}$. On aurait

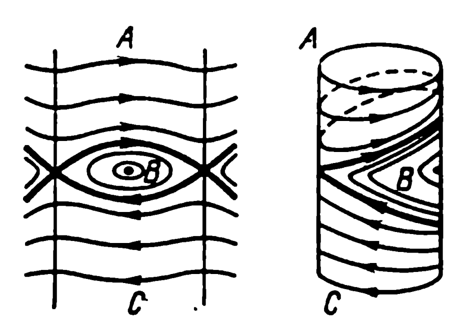


Fig. 220. Variables action-angle sur une variété symplectique.

pu introduire les variables action-angle pour un système donné sur une variété symplectique arbitraire. Nous nous limiterons à un cas simple (fig. 220).

Il est naturel de prendre pour espace des phases du pendule ($H = \frac{1}{2}p^2 - \cos q$) non pas le plan $\{(p, q)\}$ mais le cylindre de phase $\mathbf{R}^1 \times S^1$ que l'on obtient en identifiant les angles q qui diffèrent d'un multiple entier de 2π.

Les lignes de niveau critiques $H = \pm 1$ partagent le cylindre en trois régions (A), (B), (C) difféomorphes chacune au produit cartésien $\mathbf{R}^1 \times S^1$. Dans chaque région on peut introduire des variables action-angle.

Dans la région bornée (B) les trajectoires fermées représentent le balancement du pendule, dans les régions non bornées, les rotations.

R e m a r q u e 3. Comme dans le cas traité, d'une façon générale les équations $F_i = f_i$ cessent d'être indépendantes et M_f d'être une variété pour certaines valeurs f_i. A ces valeurs critiques de f

correspondent des séparatrices qui partagent l'espace des phases du problème à intégrer en régions semblables aux régions (A), (B), (C) précédentes. Dans certaines de ces régions les variétés M_f peuvent être infinies (régions (A) et (C) sur le plan $\{(p, q)\}$); les autres se décomposent en tores invariants M_f de dimension n; au voisinage d'un tel tore on peut introduire les variables action-angle.

§ 51. Moyenne

Dans ce paragraphe on démontre l'équivalence des moyennes temporelles et spatiales dans les systèmes en mouvement quasi périodique.

A. Mouvement quasi périodique. Dans les paragraphes précédents nous avons souvent rencontré le mouvement quasi périodique: figures de Lissajous, précession, nutation, rotation de la toupie, etc.

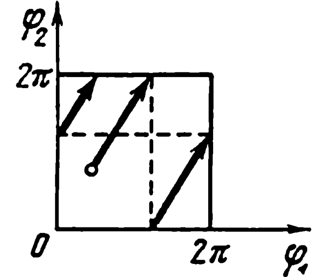


Fig. 221. Mouvement quasi périodique.

D é f i n i t i o n. Soit T^n un tore de dimension n et soient $\varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi$ des coordonnées angulaires. On appelle *mouvement quasi périodique* un groupe à un paramètre de difféomorphismes de $T^n \to T^n$ défini par les équations différentielles (fig. 221):

$$\dot{\varphi} = \omega, \quad \omega = (\omega_1, \ldots, \omega_n) = \text{const.}$$

Ces équations différentielles s'intègrent immédiatement

$$\varphi(t) = \varphi(0) + \omega t.$$

Donc sur la carte $\{\varphi\}$ les trajectoires sont des droites. On les appelle *hélices* du tore.

E x e m p l e. Soit $n = 2$. Si $\omega_1/\omega_2 = k_1/k_2$, les trajectoires sont fermées; si ω_1/ω_2 est irrationnel, les trajectoires sont partout denses sur le tore (voir § 16).

Les quantités $\omega_1, \ldots, \omega_n$ sont appelées *fréquences* du mouvement quasi périodique. Elles sont *indépendantes* par définition si elles sont linéairement indépendantes sur le corps des rationnels, i.e. si $\boldsymbol{k} \in Z^n$ *) et $(\boldsymbol{k}, \omega) = 0$, alors $\boldsymbol{k} = 0$.

B. Moyennes spatiale et temporelle. Soit $f(\varphi)$ une fonction intégrable sur le tore T^n.

*) $\boldsymbol{k} = (k_1, \ldots, k_n)$ avec k_i entiers.

Définition. On appelle *moyenne spatiale* de la fonction f sur le tore T^n le nombre

$$\overline{f} = (2\pi)^{-n} \int_0^{2\pi} \dots \int_0^{2\pi} f(\varphi)\, d\varphi_1 \dots d\varphi_n.$$

Considérons la valeur de la fonction $f(\varphi)$ sur la trajectoire $\varphi(t) = \varphi_0 + \omega t$. C'est une fonction du temps $f(\varphi_0 + \omega t)$. Etudions sa moyenne.

Définition. On appelle *moyenne temporelle* de la fonction f sur le tore T^n la fonction

$$f^*(\varphi_0) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(\varphi_0 + \omega t)\, dt$$

(qui est définie là où la limite existe)

Théorème de la moyenne. *La moyenne temporelle est partout définie et se confond avec la moyenne spatiale pourvu que la fonction f soit continue (ou à la rigueur intégrable au sens de Riemann) et les fréquences ω_i indépendantes.*

Exercice. Montrer que si les fréquences ne sont pas indépendantes, la moyenne temporelle est susceptible de ne pas se confondre partout avec la moyenne spatiale.

Corollaire 1. *Si les fréquences sont indépendantes, alors toute trajectoire $\{\varphi(t)\}$ est partout dense sur le tore T^n.*

Démonstration. Supposons le contraire. Alors le voisinage D d'un certain point du tore ne contient pas de points de la trajectoire $\varphi(t)$. Il est aisé de construire une fonction continue f nulle à l'extérieur de D et de moyenne spatiale $\overline{f} = 1$. La moyenne temporelle $f^*(\varphi_0)$ sur la trajectoire $\varphi(t)$ est égale à $0 \neq 1$. Ce qui démontre le corollaire 1.

Corollaire 2. *Si les fréquences sont indépendantes alors toute trajectoire est uniformément distribuée sur le tore T^n.*

Cela signifie que le temps de séjour de la trajectoire dans le domaine D est proportionnel à la mesure sur D.

Plus exactement, supposons que D est un domaine de T^n mesurable au sens de Jordan. Désignons par $\tau_D(T)$ le temps de séjour de la portion $0 \leqslant t \leqslant T$ de trajectoire $\varphi(t)$ à l'intérieur de D. Alors

$$\lim_{T \to \infty} \frac{\tau_D(T)}{T} = \frac{\text{mes } D}{(2\pi)^n}.$$

Démonstration. Appliquons le théorème à la fonction caractéristique f de l'ensemble D (f est intégrable au sens de Rie-

mann puisque le domaine D est mesurable au sens de Jordan). Alors $\int_0^T f(\varphi(t))\,dt = \tau_D(T)$ et $\bar{f} = (2\pi)^{-n}$ mes D et le corollaire découle immédiatement du théorème.

Corollaire. Dans la suite

$$1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2 \ldots$$

des premières décimales des nombres 2^n le chiffre 7 se rencontre $\frac{\lg 8 - \lg 7}{\lg 9 - \lg 8}$ fois plus que le chiffre 8.

Le théorème de la moyenne se rencontre implicitement dans les travaux de Laplace, Lagrange et Gauss en mécanique céleste; c'est l'un des premiers « théorèmes ergodiques ». La démonstration rigoureuse n'en a été donnée qu'en 1909 par P. Bohl, W. Sierpinski et H. Weyl dans la résolution du problème de Lagrange relatif au mouvement moyen du périhélie de la Terre. La démonstration de H. Weyl est reproduite plus bas.

C. Démonstration du théorème de la moyenne.

Lemme 1. *Le théorème est vrai pour les exponentielles* $f = e^{i(k,\,\varphi)}$, $k \in Z^n$.

D é m o n s t r a t i o n. Si $k = 0$, alors $\bar{f} = f = f^* = 1$ et le théorème est évident. Si $k \neq 0$, alors $\bar{f} = 0$. Par ailleurs

$$\int_0^T e^{i(k,\,\varphi_0+\omega t)}\,dt = e^{i(k,\,\varphi_0)}\,\frac{e^{i(k,\,\omega)T}-1}{i(k,\,\omega)}.$$

Donc la moyenne temporelle

$$\lim_{T\to\infty} \frac{e^{i(k,\,\varphi_0)}}{i(k,\,\omega)} \cdot \frac{e^{i(k,\,\omega)T}-1}{T} = 0,$$

c.q.f.d.

Lemme 2. *Le théorème est vrai pour les polynômes trigonométriques*

$$f = \sum_{|k|<N} f_k e^{i(k,\,\varphi)}.$$

D é m o n s t r a t i o n. La moyenne temporelle comme la moyenne spatiale dépendent linéairement de f, c'est pourquoi elles se confondent en vertu du lemme 1, c.q.f.d.

Lemme 3. *Soit f une fonction réelle continue (ou à la rigueur intégrable au sens de Riemann). Alors quel que soit $\varepsilon > 0$ il existe deux polynômes trigonométriques P_1 et P_2 tels que* $P_1 < f < P_2$, $\frac{1}{(2\pi)^n}\int_{T^n} (P_2 - P_1)\,d\varphi < \varepsilon$.

D é m o n s t r a t i o n. Supposons d'abord que f est continue. En vertu du théorème de Weierstrass on peut l'approcher par un polynôme trigonométrique P, $|f - P| < \frac{\varepsilon}{2}$. Les polynômes $P_1 = P - \frac{\varepsilon}{2}$ et $P_2 = P + \frac{\varepsilon}{2}$ sont les polynômes cherchés.

Si f est discontinue mais intégrable au sens de Riemann, il existe deux fonctions continues f_1 et f_2 telles que $f_1 < f < f_2$, $(2\pi)^{-n} \int (f_2 - f_1)\, d\varphi < \frac{\varepsilon}{2}$ (la figure 222 correspond à la fonction caractéristique du segment). En approchant f_1 et f_2 par les polynômes $P_1 < f_1 < f_2 < P_2$, $(2\pi)^{-n} \int (P_2 - f_2)\, d\varphi < \frac{\varepsilon}{3}$, $(2\pi)^{-n} \int (f_1 - P_1) d\varphi < < \frac{\varepsilon}{3}$, on obtient le résultat annoncé. Le lemme 3 est démontré.

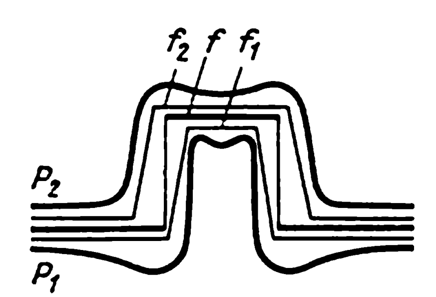


Fig. 222. Approximation de la fonction f par les polynômes trigonométriques P_1 et P_2.

Il est aisé maintenant d'achever la démonstration du théorème. Soit $\varepsilon > 0$. Alors en vertu du lemme 3 il existe des polynômes trigonométriques $P_1 < f < P_2$, $(2\pi)^{-n} \int (P_2 - P_1)\, d\varphi < \varepsilon$.

Quel que soit T on a alors

$$\frac{1}{T} \int_0^T P_1(\varphi(t))\, dt < \frac{1}{T} \int_0^T f(\varphi(t))\, dt < \frac{1}{T} \int_0^T P_2(\varphi(t))\, dt.$$

En vertu du lemme 2 pour $T > T_0(\varepsilon)$ il vient

$$\left| \overline{P}_i - \frac{1}{T} \int_0^T P_i(\varphi(t))\, dt \right| < \varepsilon \qquad (i = 1, 2).$$

Par ailleurs $\overline{P}_1 < \overline{f} < \overline{P}_2$ et $\overline{P}_2 - \overline{P}_1 < \varepsilon$. Donc $\overline{P}_2 - \overline{f} < \varepsilon$, $\overline{f} - \overline{P}_1 < \varepsilon$; donc pour $T > T_0(\varepsilon)$ on a

$$\left| \frac{1}{T} \int_0^T f(\varphi(t))\, dt - \overline{f} \right| < 2\varepsilon,$$

c.q.f.d.

E x e r c i c e. Un oscillateur de dimension deux d'énergie cinétique $T = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2$ et d'énergie potentielle $U = \frac{1}{2}x^2 + y^2$ effectue des oscilla-

tions d'amplitudes $a_x = 1$, $a_y = 1$. Trouver l'énergie cinétique moyenne temporelle.

E x e r c i c e *). Supposons que les ω_k sont indépendants, $a_k > 0$. Calculer

$$\lim_{t\to\infty} \frac{1}{t} \arg \sum_{k=1}^{3} a_k e^{i\omega_k t}.$$

Réponse. $\dfrac{\omega_1\alpha_1 + \omega_2\alpha_2 + \omega_3\alpha_3}{\pi}$, où α_1, α_2, α_3 sont les angles d'un triangle de côtés a_k (fig. 223).

D. Dégénérescences. Jusqu'ici nous n'avons étudié que les cas où les fréquences ω étaient indépendantes. Un vecteur entier $k \in Z^n$ est appelé *relation de fréquences* si $(k, \omega) = 0$.

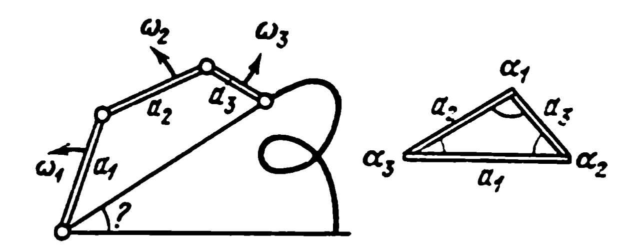


Fig. 223. Problème du mouvement moyen des périhélies des planètes.

E x e r c i c e. Montrer que l'*ensemble de toutes les relations de fréquences données* ω *forme un sous-groupe G du treillis* Z^n.

Or nous avons vu au § 49 qu'un tel sous-groupe est composé de toutes les combinaisons linéaires entières de r vecteurs indépendants k_i, $1 \leqslant r \leqslant n$. Nous dirons qu'entre les fréquences *il existe r relations* **) (*indépendantes*).

E x e r c i c e. Montrer que l'*adhérence des trajectoires* $\{\varphi(t) = \varphi_0 + \varphi t\}$ (sur T^n) *est un tore de dimension* $n - r$, *s'il existe r relations indépendantes entre les fréquences* ω; *dans ce cas le mouvement sur* T^{n-r} *est quasi périodique et ses* $n - r$ *fréquences sont indépendantes.*

Revenons maintenant au système hamiltonien intégrable décrit dans les variables action-angle I, φ par les équations

$$\dot{I} = 0, \quad \dot{\varphi} = \omega(I), \quad \text{où} \quad \omega(I) = \frac{\partial H}{\partial I}.$$

---

*) Lagrange a prouvé qu'à ce problème se ramène le mouvement moyen des périhélies des planètes. On trouvera la solution de ce problème dans les travaux de H. Weyl. L'excentricité de l'orbite de la Terre varie comme le module de la somme analogue. Les périodes glaciaires sont probablement liées aux variations de l'excentricité.

**) Démontrer que le nombre r ne dépend pas du choix des vecteurs indépendants k_i.

Dans l'espace des phases de dimension $2n$ chaque tore I = const de dimension n est invariant et le mouvement y est quasi périodique.

D é f i n i t i o n. Un système est *non dégénéré* si est non nul le déterminant

$$\det \frac{\partial \omega}{\partial I} = \det \frac{\partial^2 H}{\partial I^2}.$$

E x e r c i c e. Montrer que *si un système est non dégénéré, alors dans un voisinage d'un point quelconque il existe des mouvements quasi périodiques à n fréquences et à un nombre moindre quelconque de fréquences.*

I n d i c a t i o n. Au lieu des variables I on peut prendre pour coordonnées locales les fréquences ω en question. Les ensembles de points ω avec un nombre quelconque r de relations $(0 \leqslant r \leqslant n)$ sont partout denses dans l'espace des collections des fréquences.

Corollaire. *Si un système est non dégénéré, les tores invariants I = const sont définis de manière unique, indépendamment du choix des coordonnées action-angle I, φ dont la construction est toujours entachée d'arbitraire* *).

En effet, les tores I = const peuvent être définis comme l'adhérence des trajectoires de phase correspondant aux fréquences indépendantes ω.

Remarquons à propos que les fréquences ω seront indépendantes pour la plupart des valeurs de I.

E x e r c i c e. Montrer que la *mesure de Lebesgue de l'ensemble des I, tels que les fréquences ω (I) soient dépendantes dans un système non dégénéré, est nulle.*

I n d i c a t i o n. Montrer tout d'abord que $\text{mes}\{\omega: \exists k \neq 0, (\omega, k) = 0\} = 0$.

Au contraire, dans un système dégénéré on peut construire des coordonnées variables action-angle telles que les tores I = const n'y soient pas les mêmes. Cela est dû au fait que les adhérences des trajectoires d'un système dégénéré qui sont des tores de dimension $k < n$ peuvent être réunies de différentes façons en tores de dimension n.

E x e m p l e 1. L'oscillateur harmonique plan $\ddot{x} = -x$; $n = 2$, $k = 1$. La séparation des variables dans les coordonnées cartésiennes et dans les coordonnées polaires donne des variables action-angle différentes et des tores différents.

E x e m p l e 2. Le mouvement keplérien plan $\left(U = -\frac{1}{r}\right)$, $n = 2$, $k = 1$. Ici aussi la séparation des variables dans les coordonnées polaires et dans les coordonnées elliptiques conduit à divers I.

---

*) On peut par exemple toujours effectuer les changements $I' = I$, $\varphi' = \varphi + S_I(I)$ ou $I_1, I_2; \varphi_1, \varphi_2 \to I_1 + I_2, I_2; \varphi_1, \varphi_2 - \varphi_1$.

§ 52. Moyennisation des perturbations

On démontre ici l'invariance adiabatique de la variable action dans un système à un degré de liberté.

A. Systèmes voisins de systèmes intégrables. Nous avons étudié plus haut un grand nombre de systèmes (problèmes à une dimension, problème des deux corps, petites oscillations, mouvements d'Euler et de Lagrange d'un solide ayant un point fixe, etc.). Nous avons étudié le caractère des trajectoires de phase dans ces systèmes: nous avons vu qu'elles étaient des « hélices de tores » partout denses sur des tores invariants de l'espace des phases et que chaque trajectoire était uniformément distribuée sur ce tore.

Il ne faut pas se hâter de conclure que cette situation est typique pour les problèmes génériques. En réalité les propriétés des trajectoires des systèmes multidimensionnels peuvent être très différentes et ne pas ressembler du tout aux propriétés des mouvements quasi périodiques. En particulier, la trajectoire d'un système à n degrés de liberté peut être partout dense dans des ensembles compliqués de dimension supérieure à n de l'espace des phases de dimension $2n$; la trajectoire peut également être partout dense et uniformément distribuée sur une variété de dimension $2n - 1$ tout entière définie par l'équation $H = h$ *). De tels systèmes sont dits « non intégrables » parce qu'ils n'admettent pas d'intégrales premières univoques indépendantes de H.

L'étude de tels systèmes est loin d'être achevée; elle fait l'objet de la « théorie ergodique ».

Une méthode de résolution des systèmes non intégrables consiste à étudier des systèmes voisins des intégrables. Exemple: le mouvement des planètes autour du soleil est proche du mouvement intégrable de deux points non attractifs autour d'un centre fixe; citons encore le problème d'une toupie pesante légèrement asymétrique et des oscillations non linéaires au voisinage d'une position d'équilibre (le problème intégrable voisin est linéaire). La méthode qui va suivre est extrêmement féconde dans l'étude de tels problèmes.

B. Principe de la moyenne. Soient I, φ les variables action-angle d'un système (« non perturbé ») intégrable de hamiltonien $H_0(I)$:

$$\dot{I} = 0, \quad \dot{\varphi} = \omega(I); \quad \omega(I) = \frac{\partial H_0}{\partial I}.$$

*) Fait partie de cette classe, par exemple, le mouvement par inertie sur les variétés de courbure négative.

19—01408

Pour système voisin « perturbé » considérons le système

$$\begin{aligned} \dot{\varphi} &= \omega(I) + \varepsilon f(I, \varphi), \\ \dot{I} &= \qquad\quad \varepsilon g(I, \varphi), \end{aligned} \tag{1}$$

où $\varepsilon \ll 1$.

Laissons provisoirement de côté le système hamiltonien et considérons un système arbitraire d'équations différentielles (1) défini sur le produit cartésien $T^k \times G$ du tore $T^k = \{\varphi = (\varphi_1, \ldots, \varphi_k) \bmod 2\pi\}$ par un domaine G d'un espace $\mathbf{R}^l : G \subset \mathbf{R}^l = \{I = (I_1, \ldots, I_l)\}$. Pour $\varepsilon = 0$ le mouvement (1) est quasi périodique à fréquences en nombre $\leqslant k$ et à tores invariants de dimension k.

Le *principe de la moyenne pour le système* (1) consiste à le remplacer par un autre appelé système moyennisé:

$$\dot{J} = \varepsilon \bar{g}(J), \quad \bar{g}(J) = (2\pi)^{-k} \int_0^{2\pi} \ldots \int_0^{2\pi} g(J, \varphi)\, d\varphi_1 \ldots d\varphi_k \tag{2}$$

dans le domaine $G \subset \mathbf{R}^l = \{J = (J_1, \ldots, J_l)\}$.

On affirme que le système (2) *approche bien le système* (1).

Remarquons que ce principe n'est ni un théorème ni un axiome ni même une définition, c'est une proposition physique, i.e. une proposition vague qui, en toute rigueur, est erronée.

Il n'empêche que ces propositions sont souvent à l'origine de très importants théorèmes des mathématiques.

Le principe de la moyenne se rencontre déjà implicitement dans les travaux de Gauss (en étudiant les perturbations des planètes Gauss a proposé de distribuer la masse de chaque planète sur son orbite proportionnellement au temps et de remplacer l'attraction des planètes par l'attraction des anneaux obtenus). Cependant à ce jour on ne connaît encore aucune étude satisfaisante sur les liens existant entre les solutions des systèmes (1) et (2) dans le cas général.

En remplaçant le système (1) par le système (2) on néglige le terme $\varepsilon\tilde{g}(I, \varphi) = \varepsilon g(I, \varphi) - \varepsilon\bar{g}(I)$ du second membre. Ce terme est de l'ordre de ε comme d'ailleurs le terme restant $\varepsilon\bar{g}$. Nous allons voir sur un exemple simple que les termes $\bar{g}$ et $\tilde{g}$ n'ont pas le même rôle dans g.

Exercice. Traitons le cas $k = l = 1$,

$$\dot{\varphi} = \omega \neq 0, \quad \dot{I} = \varepsilon g(\varphi).$$

Montrer que pour $0 < t < \frac{1}{\varepsilon}$

$$|I(t) - J(t)| < c\varepsilon, \text{ où } J(t) = I(0) + \varepsilon \bar{g} t.$$

Solution.

$$I(t) - I(0) = \int_0^t \varepsilon g(\varphi_0 + \omega t)\, dt = \int_0^t \varepsilon \bar{g}\, dt +$$

$$+ \frac{\varepsilon}{\omega} \int_0^{t/\omega} \tilde{g}(\varphi)\, d\varphi = \varepsilon \bar{g} t + \frac{\varepsilon}{\omega} h\left(\frac{t}{\omega}\right),$$

où $h(\varphi) = \int_0^\varphi \tilde{g}(\varphi)\, d\varphi$ est une fonction périodique et partant bornée.

Donc les variations de I en fonction du temps se traduisent par des oscillations d'ordre ε dépendant de $\tilde{g}$ et une « évolution » systématique de vitesse $\varepsilon \bar{g}$ (fig. 224).

Le principe de la moyenne repose sur le fait que dans le cas général le mouvement du système (1) peut se décomposer en une « évolution » (2) et en petites oscillations. Dans le cas général cette décomposition est injustifiée et le principe lui-même est faux.

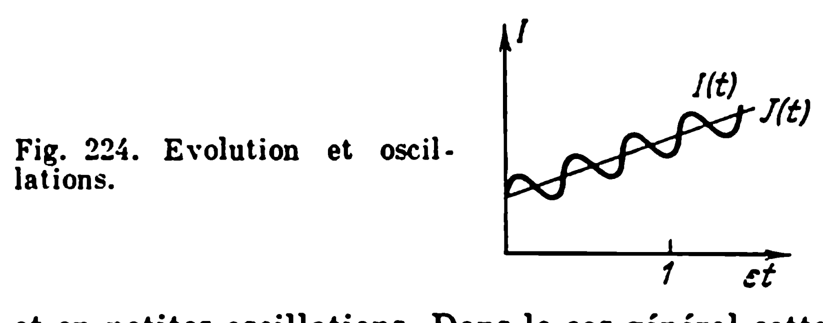


Fig. 224. Evolution et oscillations.

Néanmoins appliquons ce principe au système hamiltonien (1):

$$\dot{\varphi} = -\frac{\partial}{\partial I}(H_0(I) + \varepsilon H_1(I, \varphi)),$$

$$\dot{I} = \frac{\partial}{\partial \varphi}(H_0(I) + \varepsilon H_1(I, \varphi)).$$

Le deuxième membre du système (2) est donc

$$\bar{g} = (2\pi)^{-n} \int_0^{2\pi} \frac{\partial}{\partial \varphi} H_1(I, \varphi)\, d\varphi = 0.$$

En d'autres termes, *il n'y a pas évolution dans un système hamiltonien non dégénéré.*

Une forme de cette conclusion qui, répétons-le, est dénuée de rigueur conduit au théorème de Laplace:

Les demi grands axes des ellipses kepleriennes des planètes n'admettent pas de perturbations séculaires.

Ce qui précède suffit à nous convaincre de l'importance du principe de la moyenne; formulons maintenant un théorème qui est la traduction de ce principe dans un cas assez particulier : les oscillations à fréquence unique ($k = 1$). Ce théorème montre que l'équation centrée décrit correctement l'évolution sur un intervalle de temps élevé $\left(0 < t < \frac{1}{\varepsilon}\right)$.

C. Moyennisation d'un système à fréquence unique. Considérons le système de $l + 1$ équations différentielles

$$\left.\begin{aligned} \dot{\varphi} &= \omega(I) + \varepsilon f(I, \varphi), \\ \dot{I} &= \qquad\quad \varepsilon g(I, \varphi), \end{aligned}\right\} \begin{aligned} &\varphi \bmod 2\pi \in S^1, \\ &I \in G \subset \mathbf{R}^l, \end{aligned} \tag{1}$$

où $f(I, \varphi + 2\pi) \equiv f(I, \varphi)$, $g(I, \varphi + 2\pi) \equiv g(I, \varphi)$, et le système moyennisé de l équations

$$\dot{J} = \varepsilon \bar{g}(J), \text{ où } \bar{g}(J) = \frac{1}{2\pi} \int_0^{2\pi} g(J, \varphi)\, d\varphi. \tag{2}$$

Désignons par $I(t)$, $\varphi(t)$ la solution du système (1) qui vérifie la condition initiale $I(0)$, $\varphi(0)$ et par $J(t)$ la solution du système (2) avec la même condition initiale $J(0) = I(0)$ (fig. 225).

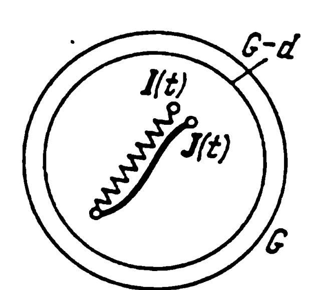


Fig. 225. Théorème de la moyenne.

Théorème. *Supposons* 1° *que les fonctions* ω, f, g *sont définies lorsque* I *varie dans un domaine borné* G, *qu'elles sont bornées dans* G *avec leurs dérivées premières et secondes* :

$$\| \omega, f, g \|_{C^2(G \times S^1)} < c_1;$$

2° *que dans le domaine* G

$$\omega(I) > c > 0;$$

3° *pour* $0 \leqslant t \leqslant \frac{1}{\varepsilon}$ *le point* $J(t)$ *est contenu dans* G *avec un voisinage de rayon* d :

$$J(t) \in G - d.$$

Alors pour un ε *suffisamment petit* $(0 < \varepsilon < \varepsilon_0)$

$$|\boldsymbol{I}(t) - \boldsymbol{J}(t)| < c_9\varepsilon \text{ pour tous les } t, \quad 0 \leqslant t \leqslant 1/\varepsilon,$$

où la constante $c_9 > 0$ *dépend de* c_1, c, d *et pas de* ε.

Certaines applications de ce théorème seront reproduites plus bas (« invariants adiabatiques »). Remarquons que l'idée maîtresse de la démonstration (à savoir le changement de variables qui a pour effet de diminuer la perturbation) est plus importante que le théorème lui-même, elle est capitale en théorie des équations différentielles ordinaires; on la rencontre déjà dans le cours élémentaire sous les traits de la méthode de variation des constantes.

D. Démonstration du théorème de la moyenne. Substituons aux variables $\boldsymbol{I}$ les nouvelles variables $\boldsymbol{P}$

$$\boldsymbol{P} = \boldsymbol{I} + \varepsilon \boldsymbol{k}(\boldsymbol{I}, \varphi), \tag{3}$$

où les fonctions $\boldsymbol{k}$ de période 2π en φ sont choisies de telle sorte que le vecteur $\boldsymbol{P}$ satisfasse à une équation différentielle plus simple.

En vertu de (1) et (3), la vitesse de variation de $\boldsymbol{P}(t)$ est égale à

$$\dot{\boldsymbol{P}} = \dot{\boldsymbol{I}} + \varepsilon \frac{\partial \boldsymbol{k}}{\partial \boldsymbol{I}} \dot{\boldsymbol{I}} + \varepsilon \frac{\partial \boldsymbol{k}}{\partial \varphi} \dot{\varphi} =$$

$$= \varepsilon \left[\boldsymbol{g}(\boldsymbol{I}, \varphi) + \frac{\partial \boldsymbol{k}}{\partial \varphi} \omega(\boldsymbol{I}) \right] + \varepsilon^2 \frac{\partial \boldsymbol{k}}{\partial \boldsymbol{I}} \boldsymbol{g} + \varepsilon^2 \frac{\partial \boldsymbol{k}}{\partial \psi} f. \tag{4}$$

Supposons que l'on puisse inverser le changement (3):

$$\boldsymbol{I} = \boldsymbol{P} + \varepsilon \boldsymbol{h}(\boldsymbol{P}, \varphi, \varepsilon) \tag{5}$$

(où les fonctions $\boldsymbol{h}$ sont de période 2π en φ).

Alors de (4) et (5) il s'ensuit que $\boldsymbol{P}(t)$ obéit au système d'équations

$$\dot{\boldsymbol{P}} = \varepsilon \left[\boldsymbol{g}(\boldsymbol{P}, \varphi) + \frac{\partial \boldsymbol{k}}{\partial \varphi} \omega(\boldsymbol{P}) \right] + \boldsymbol{R}, \tag{6}$$

où le terme résiduel $\boldsymbol{R}$ est du deuxième ordre de petitesse par rapport à ε:

$$|\boldsymbol{R}| < c_2\varepsilon^2, \quad c_2(c_1, c_3, c_4) > 0, \tag{7}$$

pourvu que

$$\|\omega\|_{C^2} < c_1, \quad \|f\|_{C^2} < c_1, \quad \|\boldsymbol{g}\|_{C^2} < c_1,$$
$$\|\boldsymbol{k}\|_{C^2} < c_3, \quad \|\boldsymbol{h}\|_{C^2} \leqslant c_4. \tag{8}$$

Essayons de choisir le changement de variables (3) de manière à annuler le membre en ε de (6). Nous obtenons pour $\boldsymbol{k}$ l'équation

$$\frac{\partial \boldsymbol{k}}{\partial \varphi} = -\frac{1}{\omega} \boldsymbol{g}.$$

D'une façon générale cette équation est insoluble dans la classe des fonctions k périodiques en φ. En effet, la valeur moyenne (sur φ) du premier membre est toujours nulle contrairement à la valeur moyenne du second.

Pour cette raison nous ne pouvons choisir k de façon à éliminer entièrement la partie de (6) qui contient ε. Cependant nous pouvons éliminer toute la partie « périodique » de g

$$\tilde{g}(P, \varphi) = g(P, \varphi) - \bar{g}(P),$$

en posant

$$k(P, \varphi) = -\int_0^{\varphi} \frac{\tilde{g}(P, \varphi)}{\omega(P)}\, d\varphi. \tag{9}$$

Nous avons défini k au moyen de la formule (9). En vertu des conditions 1 et 2 du théorème démontré la fonction k admet la majoration $\| k \|_{C^2} < c_3$, où $c_3(c_1, c) > 0$. Pour établir les inégalités (8) il reste à majorer h. Pour cela il nous faut d'abord démontrer que le changement (3) est inversible.

Fixons un nombre positif α.

Lemme. *Si ε est suffisamment petit, la restriction de l'application* (3) *)

$$I \to I + \varepsilon k, \text{ où } | k |_{C^2(G)} < c_3$$

au domaine $G - \alpha$ (composé des points de G avec un α-voisinage) est un difféomorphisme. Dans le domaine $G - 2\alpha$ le difféomorphisme inverse (5) *vérifie la majoration $\| h \|_{C^2} < c_4$, où la constante $c_4(\alpha, c_3) > 0$.*

D é m o n s t r a t i o n. La majoration cherchée découle immédiatement du théorème des fonctions implicites. La seule difficulté sera de prouver la bijectivité de l'application $I \to I + \varepsilon k$ dans le domaine $G - \alpha$.

Remarquons que la fonction k satisfait dans le domaine $G - \alpha$ une condition de Lipchitz (dans un rapport $L(\alpha, c_3)$). Soient deux points I_1, I_2 de $G - \alpha$. Pour ε suffisamment petit (notamment pour $L\varepsilon < 1$) la distance entre $\varepsilon k(I_1)$ et $\varepsilon k(I_2)$ sera inférieure à $| I_1 - I_2 |$ et $I_1 + \varepsilon k(I_1) \neq I_2 + \varepsilon k(I_2)$. Donc l'application (3) est bijective dans $G - \alpha$ et le lemme est démontré.

On déduit immédiatement que pour ε suffisamment petit toutes les majorations (8) sont valables. Donc la majoration (7) l'est également.

Comparons maintenant des systèmes d'équations différentielles pour J

$$\dot{J} = \varepsilon \bar{g}(J) \tag{2}$$

*) Quelle que soit la valeur fixe du paramètre φ.

et pour P; en vertu de (9) le dernier s'écrit

$$\dot{P} = \varepsilon \bar{g}(P) + R. \tag{6'}$$

Comme la différence entre les seconds membres est de l'ordre de $\leqslant \varepsilon^2$ (voir (7)), pendant le temps $t \leqslant 1/\varepsilon$ les solutions s'éloigne-

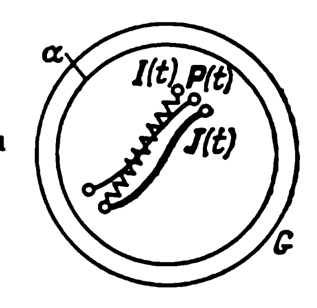


Fig. 226. Démonstration du théorème de la moyenne.

ront l'une de l'autre de $|P - J| \leqslant \varepsilon$ (fig. 226). Par ailleurs $|I - P| = \varepsilon\,|k| \leqslant \varepsilon$. Donc la différence $|I - J|$ pour $t \leqslant 1/\varepsilon$ est de l'ordre de $\leqslant \varepsilon$, c.q.f.d.

On se propose d'obtenir une meilleure majoration. Introduisons à cet effet la grandeur

$$z(t) = P(t) - J(t). \tag{10}$$

Alors de (6') et (9) il résulte

$$\dot{z} = \varepsilon(\bar{g}(P) - \bar{g}(J)) + R = \varepsilon \frac{\partial \bar{g}}{\partial P} z + R',$$

où $|R'| < c^2\varepsilon^2 + c_5\varepsilon\,|z|$ si l'intervalle (P, J) est contenu dans $G - \alpha$. Sous cette hypothèse on trouve

$$|\dot{z}| \leqslant c_6\varepsilon\,|z| + c_2\varepsilon^2 \quad (c_6 = c_5 + c_1), \quad |z(0)| \leqslant c_3\,\varepsilon. \tag{11}$$

Lemme. *Si* $|\dot{z}| \leqslant a\,|z| + b$, $|z(0)| < d$; $a, b, d, t > 0$, *alors* $|z(t)| \leqslant (d + bt)\,e^{at}$.

Démonstration. $|z(t)|$ n'est pas supérieur à la solution $y(t)$ de l'équation $\dot{y} = ay + b$, $y(0) = d$. La résolution de cette équation donne $y = Ce^{at}$, $\dot{C}e^{at} = b$, $\dot{C} = e^{-at}b$, $C(0) = d$, $C \leqslant d + bt$, c.q.f.d.

De (11) et sous l'hypothèse que l'intervalle (P, J) est contenu dans $G - \alpha$ (fig. 226)

$$|z(t)| < (c_3\varepsilon + c_2\varepsilon^2 t)\,e^{c_6\varepsilon t}.$$

D'où il résulte que pour $0 \leqslant t \leqslant 1/\varepsilon$

$$|z(t)| < c_7\,\varepsilon, \quad c_7 = (c_3 + c_2)\,e^{c_6}.$$

Nous constatons maintenant que si $\alpha = d/3$ et ε est suffisamment petit, l'intervalle $(P(t), J(t))$ $(t \leqslant 1/\varepsilon)$] sera entièrement contenu à l'intérieur de $G - \alpha$ et par conséquent

$$|P(t) - J(t)| < c_8\varepsilon \text{ pour tous les } t,\ 0 \leqslant t \leqslant 1/\varepsilon.$$

Par ailleurs $|P(t) - I(t)| < |\varepsilon k| < c_3\varepsilon$. Donc pour tous les t, $0 \leqslant t \leqslant 1/\varepsilon$,

$$|I(t) - J(t)| < c_9\varepsilon, \quad c_9 = c_8 + c_3 > 0,$$

et le théorème est démontré.

E. Invariants adiabatiques. Considérons un système hamiltonien à un degré de liberté dont la fonction de Hamilton $H(p, q; \lambda)$ dépend du paramètre λ.

Un tel système peut être le pendule

$$H = \frac{p^2}{2l^2} + lg\,\frac{q^2}{2},$$

où pour paramètre λ on peutprendre la longueur l ou l'accélération de la pesanteur g.

Supposons que le paramètre varie lentement avec le temps. Il s'avère alors qu'à la limite, lorsque la vitesse de variation du paramètre tend vers zéro, il se manifeste un remarquable phénomène asymptotique: deux grandeurs généralement indépendantes deviennent dépendantes l'une de l'autre.

Supposons par exemple que la longueur du pendule varie lentement (par rapport à ses oscillations propres). On constate que l'amplitude des oscillations devient une fonction de la longueur du pendule. Si par exemple on double très lentement la longueur du pendule, puis on la ramène à sa valeur initiale, on s'aperçoit qu'à la fin de cette opération l'amplitude des oscillations est la même qu'au départ.

Bien plus, on remarque que le rapport de l'énergie H du pendule à la fréquence ω ne change pratiquement pas sous l'effet d'une variation lente des paramètres, quoique l'énergie et la fréquence se soient fortement modifiées. De telles grandeurs sont appelées invariants adiabatiques en physique.

On conçoit aisément que l'invariance adiabatique du rapport de l'énergie du pendule à la fréquence est une affirmation à caractère physique, i.e. n'est vraie que sous certaines conditions supplémentaires. En effet, en modifiant la longueur du pendule aussi lentement que l'on veut mais en choisissant la phase d'oscillations qui assure l'allongement et la réduction de la longueur, on peut balancer le pendule (résonance paramétrique). Ayant conscience de cela les physiciens ont proposé de formuler l'invariance adiabatique comme suit: la personne qui change les paramètres du système ne doit pas voir dans quel état se trouve ce dernier (fig. 227). Définir mathématiquement cette situation est une chose pour le moins délicate et aucune tentative heureuse n'est connue à ce jour. Heureusement il existe une solution de rechange qui consiste à remplacer la non-ingérence de la personne qui modifie les paramètres dans les « affaires intérieures » du système par une condition de variation continue des paramètres, plus précisément cette variation doit être deux fois continûment différentiable.

De façon précise, soit $H(p, q; \lambda)$ une fonction fixée deux fois continûment différentiable par rapport à tous ses arguments. Posons $\lambda = \varepsilon t$ et considérons le système à paramètre lentement variable

$\lambda = \varepsilon t$:

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad \dot{q} = \frac{\partial H}{\partial p}, \quad H = H(p, q; \varepsilon t). \qquad (*)$$

Définition. On dit que la grandeur $I(p, q; \lambda)$ est un *invariant adiabatique* du système (*) si pour tout $\varkappa > 0$ il existe un $\varepsilon_0 > 0$ tel que

$$|I(p, (t), q(t); \varepsilon t) - I(p(0), q(0); 0)| < \varkappa$$

sous réserve que $\varepsilon < \varepsilon_0$, $0 < t < 1/\varepsilon$.

Toute intégrale première est visiblement invariant abiabatique.



Fig. 227. Variation adiabatique de la longueur du pendule.

Fig. 228. Invariant adiabatique d'un système de dimension un.

Il s'avère que *tout système à une dimension* (*) *possède un invariant adiabatique. Plus exactement, l'invariant adiabatique est la variable action dans le problème à coefficient constant associé.*

Supposons que les trajectoires de phase du système de hamiltonien $H(p, q; \lambda)$ sont fermées. Définissons une fonction $I(p, q; \lambda)$ comme suit. Pour λ fixe la fonction de Hamilton $H(p, q; \lambda)$ est représentée par le graphe de la fig. 228. Soit une trajectoire de phase fermée passant par un point (p, q). Elle intercepte une aire que nous désignerons par $2\pi I(p, q; \lambda)$. Sur chaque trajectoire de phase (pour λ donné) on a $I = \text{const}$. Dans I on reconnaît de toute évidence la variable d'action (voir § 50).

Théorème. *Si la fréquence* $\omega(I, \lambda)$ *du système considéré* (*) *ne s'annule pas, alors* $I(p, q; \lambda)$ *est invariant adiabatique.*

F. Démonstration de l'invariance adiabatique de l'action. Pour λ fixe dans le système (*) on peut introduire les variables action-angle I, φ moyennant une transformation canonique qui dépend de λ:

$$p, q \to I, \varphi; \ \dot{\varphi} = \omega(I, \lambda), \ \dot{I} = 0; \ \omega(I, \lambda) = \frac{\partial H^0}{\partial I},$$
$$H_0 = H_0(I, \lambda).$$

Désignons par $S(I, q; \lambda)$ la fonction génératrice (multivoque) de cette transformation:

$$p = \frac{\partial S}{\partial q}, \quad \varphi = \frac{\partial S}{\partial I}.$$

Soit maintenant $\lambda = \varepsilon t$. Comme on passe des variables p, q aux variables I, φ par une transformation canonique *fonction du temps* cette fois, les équations du mouvement dans les nouvelles variables I, φ prendront la forme de Hamilton mais leur hamiltonien sera (voir § 45, A):

$$K = H_0 + \frac{\partial S}{\partial t} = H_0 + \varepsilon \frac{\partial S}{\partial \lambda}.$$

Exercice. Montrer que $\frac{\partial}{\partial \lambda} S\,(I,\ q;\ \lambda)$ est une fonction univoque dans le plan de phase.

Indication. La non-univocité de S se ramène à l'addition de multiples de $2\pi I$.

Nous obtenons les équations de mouvement sous la forme

$$\begin{aligned} \dot\varphi &= \omega\,(I,\ \lambda) + \varepsilon f\,(I,\ \varphi;\ \lambda), \quad f = -\partial^2 S/\partial I\,\partial\lambda,\\ \dot I &= \qquad\qquad\quad \varepsilon g\,(I,\ \varphi;\ \lambda), \quad g = \partial^2 S/\partial\varphi\,\partial\lambda,\\ \dot\lambda &= \qquad\qquad\quad \varepsilon. \end{aligned}$$

Comme $\omega \neq 0$, on peut utiliser le théorème de la moyenne (page 292). Le système moyennisé s'écrit

$$\dot J = \varepsilon \bar g, \quad \dot\Lambda = \varepsilon.$$

Or $g = \frac{\partial}{\partial\varphi}\frac{\partial S}{\partial\lambda}$ et $\frac{\partial S}{\partial\lambda}$ est une fonction univoque sur le cercle $I =$ $=$ const. Donc $\bar g = (2\pi)^{-1} \int g\,d\varphi = 0$ et J est invariant dans le système centré: $J\,(t) = J\,(0)$.

En vertu du théorème de la moyenne

$$|\,I\,(t) - I\,(0)\,| < c\varepsilon$$

pour tous les t, $0 \leqslant t \leqslant 1/\varepsilon$, c.q.f.d.

Exemple. Dans le cas de l'oscillateur harmonique (voir fig. 217)

$$H = \frac{a^2}{2}p^2 + \frac{b^2}{2}q^2, \quad I = \frac{1}{2\pi}\pi\frac{\sqrt{2h}}{a}\frac{\sqrt{2h}}{b} = \frac{h}{\omega}, \quad \omega = ab,$$

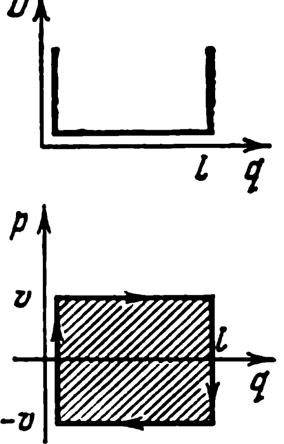
i.e. l'invariant adiabatique est le quotient de l'énergie par la fréquence.

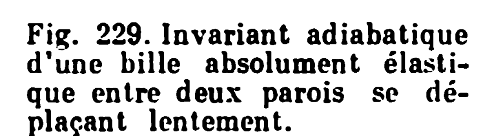
Exercice. On double lentement la longueur d'un pendule ($l = l_0\,(1 + \varepsilon t)$, $0 \leqslant t \leqslant 1/\varepsilon$). Etudier les variations de l'angle d'écart de l'amplitude q_{max}.

Solution. $I = \frac{1}{2} l^{3/2} g^{1/2} q_{max}^2$, c'est pourquoi

$$q_{max}\,(t) = q_{max}\,(0)\left(\frac{l\,(0)}{l\,(t)}\right)^{3/4}.$$

Comme deuxième exemple considérons le mouvement d'une bille absolument élastique de masse unité entre deux parois absolument élastiques dont la distance l varie lentement (fig. 229).




Fig. 229. Invariant adiabatique d'une bille absolument élastique entre deux parois se déplaçant lentement.

On peut s'imaginer que le point se déplace dans une fosse de potentiel rectangulaire de profondeur infinie et que les trajectoires de phase sont des rectangles d'aire $2vl$, où v est la vitesse de la bille. Dans ce cas l'invariant adiabatique est le produit vl de la vitesse de la bille par la distance des parois *).

Si donc on réduit de moitié la distance des parois, la vitesse de la bille double, et si l'on éloigne les parois l'une de l'autre, la vitesse diminue.

---

*) Ceci ne découle pas formellement du théorème démontré, puisqu'il y est question de systèmes différentiables, sans chocs. La démonstration de l'invariance adiabatique de vl dans ce système est un problème élémentaire très instructif.

APPENDICE 1

COURBURE RIEMANNIENNE

En roulant une feuille de papier on peut construire un cône ou un cylindre, mais on ne peut obtenir aucun morceau de sphère sans plier, distendre ou déchirer cette feuille. La raison est que les « géométries internes » de ces surfaces sont différentes : aucune région de la sphère ne peut être appliquée isométriquement sur le plan.

Les métriques riemanniennes se caractérisent par une grandeur invariante appelée courbure riemannienne. La courbure riemannienne du plan est nulle, celle d'une sphère de rayon R est égale à R^{-2}. Si une variété riemannienne s'applique isométriquement sur une autre, les courbures riemanniennes sont égales aux points correspondants. Par exemple, le cylindre ou le cône étant localement isomorphes au plan, leur courbure riemannienne est nulle en tout point. Donc aucune région du cône ou du cylindre ne peut être isométriquement appliquée sur la sphère.

La courbure riemannienne exerce une influence très notable sur le comportement des géodésiques sur une variété, i.e. sur le mouvement du système dynamique correspondant. Si la courbure riemannienne de la variété est positive (comme dans la sphère ou l'ellipsoïde), les géodésiques voisines oscillent l'une près de l'autre dans la plupart des cas, si au contraire elle est négative (comme dans l'hyperboloïde à une nappe), elles s'éloignent très vite l'une de l'autre.

Dans cet appendice on définit la courbure riemannienne et on traite brièvement les propriétés des géodésiques sur des variétés de courbure négative. Pour de plus amples renseignements sur la courbure riemannienne le lecteur pourra se référer à l'ouvrage de J. Milnor, *Morse Theory*, Princeton, 1963, et sur les géodésiques sur les variétés à courbure négative à l'ouvrage de D. Anossov, *Flots géodésiques sur les variétés de courbure négative* (en russe). Travaux de l'institut Steklov, 1967.

A. Transport parallèle sur une surface. La définition de la courbure riemannienne est basée sur la construction d'un transport

parallèle de vecteurs le long de courbes situées sur une variété riemannienne.

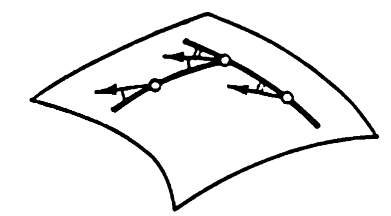
Commençons par l'exemple où la variété donnée est de dimension deux, i.e. est une surface, et la courbe donnée une géodésique de cette surface.

Le *transport parallèle d'un vecteur tangent à une surface le long d'une géodésique* se définit comme suit: le point d'application du vecteur se déplace sur la géodésique de telle façon que le module de ce vecteur et l'angle qu'il fait avec cette géodésique restent constants.

Le transport en l'extrémité d'une géodésique de tous les vecteurs tangents à la surface, à l'origine de cette géodésique, définit une application du plan tangent à l'origine dans le plan tangent à l'extrémité. Cette application est linéaire et isométrique.

Définissons maintenant le *transport parallèle d'un vecteur sur une surface le long d'une ligne polygonale* constituée d'arcs de géodé-

Fig. 230. Transport parallèle le long d'une géodésique polygonale.



siques (fig. 230). Pour transporter un vecteur le long d'une ligne polygonale nous le déplaçons successivement d'un sommet à l'autre le long des arcs de géodésiques.

E x e r c i c e. Transporter d'un tour de périmètre un vecteur tangent à une sphère au sommet d'un triangle sphérique rectangle.

Réponse. Ce transport fera tourner de 90° le plan tangent à la sphère au sommet envisagé.

Enfin, le *transport parallèle d'un vecteur le long d'une courbe différentiable quelconque tracée sur une surface* se définit à l'aide d'un passage à la limite dans lequel la courbe est approchée par des lignes polygonales constituées d'arcs de géodésiques.

E x e r c i c e. Transporter un vecteur dirigé vers le nord, de Léningrad à Léningrad (à la latitude 60°), le long du 60-ème parallèle vers l'est.

Réponse. Le vecteur subira une rotation de $2\pi\,(1 - \sin\lambda)$, i.e. environ de 50° vers l'ouest. Donc l'angle de rotation est proportionnel à l'aire de la calotte sphérique interceptée par notre parallèle; quant au sens de rotation, il est le même que le sens suivant lequel on contourne le pôle nord pendant le transport.

I n d i c a t i o n. Il suffit de transporter le vecteur le long d'un même cercle sur le cône engendré par les tangentes en tous les points du parallèle dirigées vers le nord (fig. 231). Ce cône peut être déployé sur un plan; par suite, le transport parallèle sur le cône se ramène à un transport parallèle usuel sur le plan.

E x e m p l e. Soit le demi-plan $y > 0$ du plan de la variable complexe $z = x + iy$ de métrique

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

Un calcul immédiat nous montre que les géodésiques de cette variété rieman-

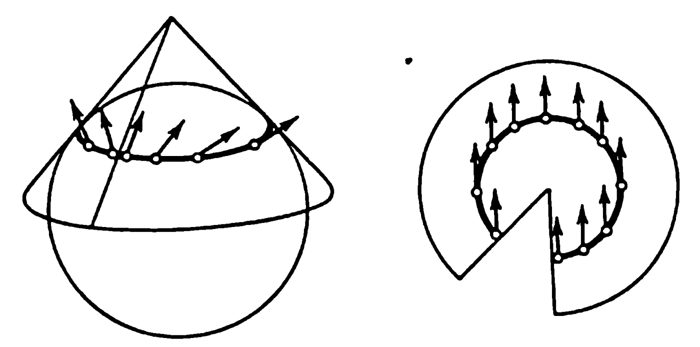


Fig. 231. Transport parallèle sur la sphère.

nienne de dimension 2 sont des cercles et des droites perpendiculaires à l'axe x. Les transformations homographiques à coefficients réels

$$z \mapsto \frac{az+b}{cz+d}$$

sont des transformations isométriques de notre variété, qui s'appelle *plan de Lobatchevski.*

E x e r c i c e. Transporter le vecteur directeur de l'axe imaginaire du point $z = i$ au point $z = t + i$ le long d'une horizontale ($dy = 0$) (fig. 232).

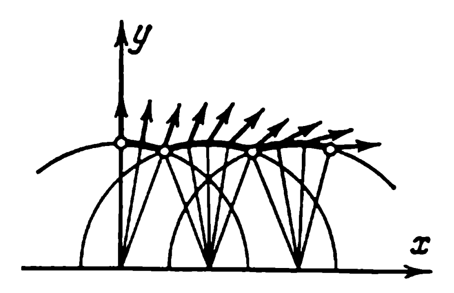


Fig. 232. Transport parallèle sur le plan de Lobatchevski.

Réponse. Lorsque le vecteur se déplace de i à $t + i$, il subit une rotation de t radians dans le sens qui va du vecteur unitaire de l'axe y à celui de l'axe x.

B. Forme de courbure. Maintenant nous sommes en mesure de définir la courbure riemannienne d'une variété riemannienne de dimension deux (i.e. une surface) en chacun de ses points. A cet effet faisons choix au voisinage du point considéré d'une orientation de la surface et soit un transport parallèle de vecteurs le long du bord d'un *petit* domaine D de cette surface. Il est clair que le résultat de ce transport est une rotation d'un angle *petit*, que nous désignerons par φ (G) (l'orientation de cet angle est fixée par l'orientation de la surface).

Si le domaine D est composé de deux régions D_1 et D_2, le transport le long du bord de D se ramène au transport le long du bord de D_1, puis de D_2. Donc

$$\varphi(D) = \varphi(D_1) + \varphi(D_2),$$

i.e. l'angle φ est une fonction additive du domaine. Si l'on modifie le sens de parcours du bord de D, l'angle φ change de signe. Il est donc naturel de représenter $\varphi(D)$ comme l'intégrale d'une 2-forme adéquate étendue à D. Une telle 2-forme existe réellement; on l'appelle *forme de courbure* et on la désigne par Ω. La forme de courbure Ω sera donc définie par la relation

$$\varphi(D) = \int_D \Omega. \tag{1}$$

La valeur de la forme Ω sur un couple ξ, η de vecteurs tangents de TM_x se calcule comme suit. Identifions un voisinage du point 0 de TM_x à un voisinage du point x sur M (par exemple au moyen de coordonnées locales). Nous pouvons alors construire sur M un parallélogramme P_ε sur des vecteurs $\varepsilon\xi$, $\varepsilon\eta$, à la rigueur pour des ε suffisamment petits.

La forme de courbure prend sur ces vecteurs la valeur

$$\Omega(\xi, \eta) = \lim_{\varepsilon\to 0} \frac{\varphi(P_\varepsilon)}{\varepsilon^2}. \tag{2}$$

En d'autres termes, *la forme de courbure prend sur un couple de vecteurs tangents infiniment petits une valeur égale à l'angle de rotation sous l'effet du transport le long du parallélogramme infiniment petit construit sur ces vecteurs.*

E x e r c i c e. Trouver la forme de courbure sur le plan, sur une sphère de rayon R et sur le plan de Lobatchevski.

Réponse. $\Omega = 0$, $\Omega = R^{-2}dS$, $\Omega = -dS$, où la 2-forme dS est l'élément de surface de notre surface orientée.

E x e r c i c e. Montrer que la fonction définie par la formule (2) est une 2-forme différentielle ne dépendant pas de la construction et que la rotation subie par le vecteur pendant le transport le long du bord du domaine orienté fini D s'exprime en fonction de cette forme au moyen de la formule (1).

E x e r c i c e. Montrer que l'intégrale de la forme de courbure étendue à une surface convexe quelconque d'un espace euclidien de dimension trois est égale à 4π.

C. Courbure riemannienne d'une surface. Remarquons maintenant que toute 2-forme différentielle sur une variété riemannienne M de dimension deux peut être écrite sous la forme $\rho\, dS$, où dS est l'élément de surface orientée et ρ une fonction numérique univoquement définie par le choix de la métrique et l'orientation.

En particulier la forme de courbure peut s'écrire

$$\Omega = K\, dS,$$

où $K: M \to \mathbf{R}$ est une fonction différentiable sur M et dS l'élément de surface.

La valeur prise par la fonction K au point x est la *courbure riemannienne de la surface* au point x.

E x e r c i c e. Calculer la courbure riemannienne du plan euclidien, d'une sphère de rayon R et du plan de Lobatchevski.

Réponse. $K = 0$, $K = R^{-2}$, $K = -1$.

E x e r c i c e. Montrer que la courbure riemannienne dépend non pas de l'orientation de la variété, mais seulement de sa métrique.

I n d i c a t i o n. Lorsqu'on modifie l'orientation, les 2-formes Ω et dS changent de signe simultanément.

E x e r c i c e. Montrer que dans l'espace euclidien ordinaire la courbure riemannienne des surfaces est en chaque point égale au produit des grandeurs inverses des principaux rayons de courbure (avec le signe moins si les centres sont situés de part et d'autre de la surface).

Remarquons que le signe de la courbure d'une variété en un point ne dépend pas de l'orientation de cette variété; on peut en effet définir ce signe sans faire intervenir l'orientation.

Plus exactement, sur une variété de courbure *positive* le transport parallèle le long du bord d'un petit domaine fait subir au vecteur une rotation autour de son point d'application *dans le sens de parcours du domaine*; sur une variété de courbure *négative* la rotation a lieu dans le sens *inverse*.

Remarquons par ailleurs que la courbure en un point est définie uniquement par la métrique au voisinage de ce point et pour cette raison se conserve sous l'effet d'une déformation isométrique: les surfaces isométriques possèdent des courbures égales aux points correspondants. Aussi la courbure riemannienne est-elle également appelée courbure intrinsèque.

Les formules de calcul de la courbure en fonction des constantes de la métrique dans un système de coordonnées quelconques renferment les dérivées secondes de la métrique et sont assez compliquées (voir exercices du point F plus bas).

D. Transport parallèle multidimensionnel. La réalisation du transport parallèle sur les variétés riemanniennes de dimension supérieure à deux est plus compliquée. En effet, dès la dimension trois la condition d'invariance de l'angle que le vecteur transporté fait avec la géodésique ne suffit pas à en définir le sens. Plus exactement, ce vecteur peut encore tourner autour de la direction de la géodésique en faisant avec elle un angle constant.

Pour préciser la construction du transport parallèle le long d'une géodésique il faut choisir un plan passant par la tangente

à la géodésique et contenant le vecteur transporté. Ce choix s'effectue de la manière suivante qui (hélas!) est assez compliquée.

A l'origine de la géodésique le plan cherché existe: c'est le plan engendré par le vecteur transporté et le vecteur directeur de la géodésique. Considérons toutes les géodésiques issues du point initial suivant les directions contenues dans ce plan. De telles géodésiques forment une surface différentiable (au voisinage du point initial) qui contient la géodésique le long de laquelle nous avons précisément l'intention de transporter le vecteur (fig. 233).

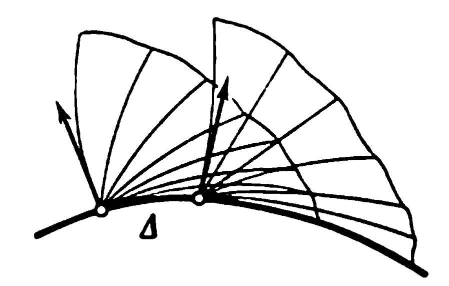


Fig. 233. Transport parallèle dans l'espace.

Considérons un autre point de cette géodésique situé à une distance Δ petite du point initial. Le plan tangent à la surface que nous venons de décrire au nouveau point renferme la direction de la géodésique en ce point. Prenons ce nouveau point pour origine et utilisons le plan tangent en ce point pour construire une nouvelle surface (engendrée par le faisceau de géodésiques issues du nouveau point). Cette surface contient la géodésique initiale. Déplaçons-nous encore d'une distance Δ le long de la géodésique et répétons cette procédure.

En un nombre fini de pas on atteint n'importe quel point de la géodésique initiale. Ce faisant on définit en chaque point un plan tangent contenant la direction de la géodésique. Ce plan dépend du pas Δ. Si $\Delta \to 0$, la famille de plans tangents obtenue tend (le calcul le montre) vers une limite définie.

Ainsi le long de notre géodésique on définit un champ de 2-plans tangents contenant la direction de la géodésique et défini intrinsèquement par la métrique de la variété.

Maintenant le transport parallèle de notre vecteur le long d'une géodésique se définit comme dans le cas bidimensionnel: ce vecteur doit rester dans les plans indiqués, conserver son module et l'angle qu'il fait avec la direction de la géodésique. Le transport parallèle le long d'une géodésique quelconque se définit par approximation par des lignes polygonales comme dans le cas bidimensionnel.

Exercice. Montrer que le transport parallèle de vecteurs d'un point d'une variété riemannienne dans un autre le long d'un chemin fixe est un opérateur isométrique linéaire de l'espace tangent au premier point dans l'espace tangent au second.

20—01408

Exercice. Transporter un vecteur arbitraire parallèlement le long de la courbe

$$x_1 = t, \quad x_2 = 0, \quad y = 1 \quad (0 \leqslant t \leqslant \tau)$$

d'un espace de Lobatchevski muni de la métrique

$$ds^2 = \frac{dx_1^2 + dx_2^2 + dy^2}{y^2}.$$

Réponse. Les vecteurs directeurs des axes x_1 et y tournent dans le plan qu'ils engendrent d'un angle τ dans le sens qui va de l'axe y à l'axe x_1; quant au vecteur directeur de x_2, il subit un déplacement parallèle au sens de la métrique euclidienne.

E. Tenseur de courbure. Voyons maintenant, comme nous l'avons fait dans le plan, le transport parallèle le long d'un petit lacet d'une variété riemannienne.

Le transport parallèle le long de ce lacet ramène les vecteurs à l'espace tangent initial. L'application obtenue de l'espace tangent dans lui-même est une petite rotation (i.e. une transformation orthogonale proche de l'identité).

Dans le plan nous avons caractérisé cette rotation par un seul nombre: l'angle de rotation φ. Dans le cas multidimensionnel le rôle de φ est assumé par un opérateur antisymétrique.

Plus exactement, tout opérateur orthogonal A voisin de l'opérateur identique s'écrit de manière unique sous la forme

$$A = e^{\Phi} = E + \Phi + \frac{\Phi^2}{2!} + \ldots,$$

où Φ est un petit opérateur antisymétrique.

Exercice. Calculer Φ si A est une rotation du plan d'angle φ petit.

Réponse. $A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, $\Phi = \begin{pmatrix} 0 & \varphi \\ -\varphi & 0 \end{pmatrix}$.

Contrairement au cas bidimensionnel la fonction du chemin Φ en général n'est pas une fonction additive (puisque tout groupe orthogonal d'un espace de dimension n n'est pas commutatif pour $n > 2$). Néanmoins nous pouvons construire à l'aide de Φ une forme de courbure décrivant « une rotation infiniment petite sous l'effet du transport le long d'un parallélogramme infiniment petit » en procédant comme dans le plan, i.e. à l'aide de la formule (2).

Soient donc ξ, η des vecteurs de TM_x. Construisons sur M un petit parallélogramme curviligne P_ε. (Les côtés du parallélogramme P_ε se déduisent des vecteurs $\varepsilon\xi$, $\varepsilon\eta$ de l'espace tangent TM_x par identification, à l'aide des coordonnées locales, d'un voisinage du zéro de TM_x avec un voisinage du point x sur M.) Considérons le transport parallèle le long des côtés du parallélogramme P_ε (on part de ξ).

Le transport se traduira par une transformation orthogonale de l'espace TM_x, voisine de la transformation identique (elle s'en

distingue d'une valeur de l'ordre de ε^2), qui s'écrit

$$A_\varepsilon(\xi, \eta) = E + \varepsilon^2\Omega + o(\varepsilon^2),$$

où Ω est un opérateur antisymétrique qui dépend de ξ et η.

Donc nous pouvons définir une fonction Ω du couple de vecteurs ξ, η de l'espace tangent au point x à valeurs dans l'espace des opérateurs antisymétriques sur TM_x au moyen de la formule

$$\Omega(\xi, \eta) = \lim_{\varepsilon\to 0} \frac{A_\varepsilon(\xi, \eta) - E}{\varepsilon^2}.$$

Exercice. Montrer que la fonction Ω est une 2-forme différentielle (à valeurs dans l'espace des opérateurs antisymétriques sur TM_x) et ne dépend pas du choix des coordonnées qui nous ont servi à identifier TM_x à M.)

La forme Ω est appelée *tenseur de courbure* de la variété riemannienne. Le tenseur de courbure décrit donc une rotation infinitésimale de l'espace tangent sous l'effet d'un transport parallèle le long d'un parallélogramme infiniment petit.

F. Courbure sectionnelle. Considérons un 2-plan L dans l'espace tangent à une variété riemannienne en l'un quelconque de ses points. Menons à partir de ce point les géodésiques dans toutes les sections du plan L. Ces géodésiques forment au voisinage du point envisagé une surface différentiable. La surface construite est plongée dans la variété riemannienne dont elle hérite la métrique riemannienne.

La *courbure de la variété riemannienne M dans une section du 2-plan L dans l'espace TM_x* est par définition la courbure riemannienne décrite plus haut.

Exercice. Trouver la courbure sectionnelle d'une sphère ordinaire de rayon R et du plan de Lobatchevski.

Réponse. R^{-2}, -1.

D'une façon générale les courbures sectionnelles d'une variété riemannienne ne prennent pas la même valeur en un même point. La formule (3) ci-dessous nous indique comment elles dépendent de la section.

Théorème. *La courbure d'une variété riemannienne dans une section définie par un couple de vecteurs orthogonaux unitaires s'exprime en fonction du tenseur de courbure Ω par la formule*

$$K = \langle \Omega(\xi, \eta)\,\xi, \eta \rangle, \tag{3}$$

où les chevrons désignent le produit scalaire qui définit la métrique riemannienne.

La démonstration se fait par comparaison des définitions du tenseur de courbure et de la courbure dans une section. Nous omettrons cette démonstration. Au besoin on pourra prendre la formule (3) pour définition de la courbure K.

G. Différentiation covariante. Au transport parallèle le long des courbes d'une variété riemannienne est lié un calcul différentiel spécial appelé différentiation covariante ou connexité riemannienne et se définissant comme suit.

Soient ξ un vecteur de TM_x, v un champ de vecteurs défini sur M au voisinage de x. La dérivée covariante du champ v suivant la direction du vecteur ξ se définit à l'aide d'une courbe quelconque issue du point x à la vitesse ξ.

En se déplaçant sur cette courbe, au bout d'un temps t on se retrouve en un point $x(t)$. Prenons le vecteur du champ v appliqué en ce point et transportons-le parallèlement en arrière le long de la courbe jusqu'au point initial x. Nous obtenons un vecteur fonction du temps t dans l'espace TM_x. Pour $t = 0$ c'est le vecteur $v(x)$, pour les autres valeurs de t il varie en fonction de la non-colinéarité des vecteurs du champ v le long de notre courbe de direction ξ.

Considérons la dérivée du vecteur obtenu par rapport à t pour $t = 0$. Cette dérivée est un vecteur de l'espace tangent TM_x. On l'appelle *dérivée covariante du champ v par rapport à* ξ et on la note $\nabla_\xi v$.

On vérifie sans peine que le vecteur $\nabla_\xi v$ ne dépend pas du choix de la courbe qui a participé à sa définition, mais seulement du vecteur ξ et du champ v.

Exercice 1. Etablir les propriétés suivantes de la différentiation covariante:

1. $\nabla_\xi v$ est une fonction bilinéaire de ξ et v.
2. $\nabla_\xi fv = (L_\xi f)\, v + f(x)\nabla_\xi v$, où f est une fonction différentiable, $L_\xi f$ la dérivée de la fonction f suivant la direction du vecteur ξ de TM_x.
3. $L_\xi \langle v, w\rangle = \langle \nabla_\xi v, w(x)\rangle + \langle v(x), \nabla_\xi w\rangle$.
4. $\nabla_{v(x)} w - \nabla_{w(x)} v = [w, v](x)$ (où $L_{[w, v]} = L_v L_w - L_w L_v$).

Exercice 2. Montrer que le tenseur de courbure s'exprime en fonction des dérivées covariantes par la formule

$$\Omega(\xi_0, \eta_0)\,\zeta_0 = -\nabla_\xi \nabla_\eta \zeta + \nabla_\eta \nabla_\xi \zeta + \nabla_{[\eta, \xi]} \zeta,$$

où ξ, η, ζ sont des champs de vecteurs quelconques qui prennent les valeurs ξ_0, η_0, ζ_0 au point considéré.

Exercice 3. Montrer que le tenseur de courbure vérifie les identités suivantes:

$$\Omega(\xi, \eta)\,\zeta + \Omega(\eta, \zeta)\,\xi + \Omega(\zeta, \xi)\,\eta = 0,$$

$$\langle \Omega(\xi, \eta)\,\alpha, \beta\rangle = \langle \Omega(\alpha, \beta)\,\xi, \eta\rangle.$$

Exercice 4. Supposons qu'une métrique riemannienne est définie dans les coordonnées locales $x_1, \ldots, x_n$ par la matrice symétrique g_{ij}:

$$ds^2 = \sum g_{ij}\, dx_i\, dx_j.$$

Désignons par $e_1, \ldots, e_n$ des champs de vecteurs arithmétiques (de sorte que la dérivée suivant la direction e_i est $\partial_i = \partial/\partial x_i$). Les dérivées covariantes

peuvent alors être calculées à l'aide des formules de l'exercice 1 et des suivantes:

$$\nabla_{e_i} e_j = \sum_k \Gamma^k_{ji} e_k, \quad \Gamma^k_{ij} = \sum_l \frac{1}{2} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}) g^{lk},$$

où (g^{lk}) est la matrice inverse de (g_{kl}).

En se servant de l'expression du tenseur de courbure de l'exercice (2) on obtient explicitement le tenseur de courbure. Les nombres $R_{ijkl} = \langle \Omega (e_i, e_j), e_k, e_l \rangle$ sont appelés composantes du tenseur de courbure.

H. Equation de Jacobi. La courbure d'une variété riemannienne est étroitement liée au comportement de ses géodésiques. Considérons en particulier une géodésique issue d'un point quelconque suivant une section quelconque et modifions légèrement les conditions initiales, i.e. le point initial et la vitesse initiale. Les nouvelles conditions initiales définissent une nouvelle géodésique qui diffère peu de la précédente. Pour étudier cet écart il est utile de linéariser l'équation différentielle des géodésiques au voisinage de la géodésique initiale.

L'équation différentielle linéaire du second ordre ainsi obtenue (« équation aux variations » pour l'équation des géodésiques) s'appelle *équation de Jacobi*. Il est plus commode de l'écrire en fonction des dérivées covariantes et du tenseur de courbure.

Désignons par $x(t)$ un point se déplaçant sur une géodésique de la variété M à une vitesse constante $v(t) \in TM_{x(t)}$.

Si la condition initiale est différentiable par rapport à un paramètre α, la géodésique le sera aussi. Considérons un mouvement correspondant à une valeur du paramètre α. Désignons par $x(t, \alpha) \in$ $\in M$ la position occupée sur la géodésique correspondante par le point au bout d'un temps t. Nous conviendrons que la géodésique initiale correspond à la valeur nulle du paramètre de sorte que $x(t, 0) = x(t)$.

On appelle *champ de vecteurs de la variation d'une géodésique* la dérivée de la fonction $x(t, \alpha)$ par rapport à α pour $\alpha = 0$; la valeur de ce champ au point $x(t)$ est égale à

$$\left.\frac{d}{d\alpha}\right|_{\alpha=0} x(t, \alpha) = \xi(t) \in TM_{x(t)}.$$

Pour écrire l'équation aux variations définissons la dérivée covariante par rapport à t d'un champ de vecteurs $\zeta(t)$ donné sur la géodésique $x(t)$. Cette définition consiste à choisir un vecteur $\zeta(t+h)$, à le transporter parallèlement du point $x(t+h)$ au point $x(t)$ le long d'une géodésique et ensuite dériver le vecteur obtenu dans l'espace tangent $TM_{x(t)}$ par rapport à h pour $h = 0$. On obtient ainsi un vecteur de l'espace tangent $TM_{x(t)}$ que l'on appelle dérivée covariante du champ $\zeta(t)$ par rapport à t et que l'on note $D\zeta/Dt$.

Théorème. *Le champ de vecteurs de la variation d'une géodésique obéit à l'équation différentielle du second ordre*

$$\frac{D^2\xi}{Dt^2} = -\Omega(v, \xi)v, \tag{4}$$

où Ω est le tenseur de courbure, $v = v(t)$ le vecteur vitesse du mouvement le long de la géodésique initiale.

Inversement, toute solution de l'équation différentielle (4) *est champ de vecteurs de la variation de la géodésique initiale.*

L'équation (4) s'appelle équation de Jacobi.

Exercice. Démontrer le théorème énoncé.

Exercice. Soient M une surface, $y(t)$ la composante d'un vecteur $\xi(t)$ suivant la normale à la géodésique considérée; on suppose que la longueur du vecteur $v(t)$ est égale à l'unité. Montrer que y obéit à l'équation différentielle

$$\ddot{y} = -Ky, \tag{5}$$

où $K = K(t)$ est la courbure riemannienne de la surface au point $x(t)$.

Exercice. Se servir de l'équation (5) pour comparer le comportement des géodésiques proches de la géodésique donnée sur la sphère ($K = +R^2$) et sur le plan de Lobatchevski ($K = -1$).

I. Discussion de l'équation de Jacobi. Lorsqu'on étudie des équations aux variations, il est utile d'éliminer les variations triviales qui consistent à modifier l'origine du temps et la grandeur de la vitesse initiale du mouvement. A cet effet décomposons le vecteur variation ξ suivant des composantes parallèle et perpendiculaire au vecteur vitesse v. Alors en vertu de $\Omega(v, v) = 0$ et de l'antisymétrie de l'opérateur $\Omega(v, \xi)$ nous obtenons pour composante normale une solution de l'équation de Jacobi et pour composante parallèle une solution de l'équation $D^2\xi/Dt^2 = 0$.

Remarquons maintenant que l'équation de Jacobi pour la composante normale peut s'écrire sous la forme de l'« équation de Newton »:

$$\frac{D^2\xi}{Dt^2} = -\operatorname{grad} U,$$

où la forme quadratique U en ξ s'exprime en fonction du tenseur de courbure et est proportionnelle à la courbure K dans une section du plan (ξ, v):

$$U(\xi) = \frac{1}{2}\langle\Omega(v, \xi)v, \xi\rangle = \frac{1}{2}K\langle\xi, \xi\rangle\langle v, v\rangle.$$

Donc la *composante normale du vecteur variation d'une géodésique parcourue à une vitesse unité obéit à l'équation d'un oscillateur linéaire (non autonome) dont l'énergie potentielle est égale au demi-produit de la courbure dans une section du plan engendré par le vecteur vitesse*

et le vecteur variation, par le carré de la composante normale de la variation.

En particulier, considérons le cas où la courbure est *négative* dans toutes les sections contenant le vecteur vitesse de la géodésique (fig. 234). Alors l'écart entre la géodésique considérée et les géodésiques voisines suivant la normale obéit à l'équation d'un oscillateur dont l'énergie potentielle (qui ne dépend pas du temps) est

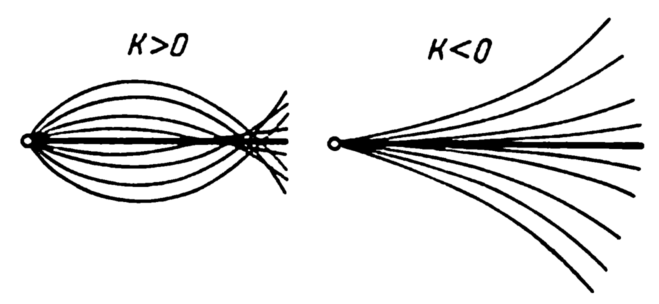


Fig. 234. Géodésiques voisines sur des variétés de courbures positive et négative.

définie négative. Donc la composante normale de l'écart des géodésiques voisines se conduit comme l'écart d'une bille située près du sommet d'une pente par rapport à ce sommet. La position d'équilibre de la bille est instable au sommet de la pente. Cela veut dire que les *géodésiques voisines de la géodésique donnée s'éloigneront exponentiellement d'elle.*

Si l'énergie potentielle de l'équation de Newton obtenue ne dépendait pas du temps, notre conclusion serait parfaitement rigoureuse. Bien plus, admettons que les courbures dans diverses sections contenant v sont comprises dans l'intervalle

$$-a^2 \leqslant K \leqslant -b^2, \text{ où } 0 < b < a.$$

Alors les solutions de l'équation de Jacobi pour les écarts normaux seront des combinaisons linéaires d'exponentielles d'exposants $\pm\lambda_i$, où les λ_i sont des nombres positifs compris entre a et b.

Donc toute solution de l'équation de Jacobi croît au moins aussi vite que $e^{b|t|}$ soit que $t \to +\infty$, soit que $t \to -\infty$; en outre la plupart des solutions croissent encore plus vite, à la vitesse $e^{a|t|}$.

L instabilité de la position d'équilibre pour une énergie potentielle définie négative est intuitive dans le cas non autonome. On peut néanmoins la prouver par comparaison avec un système autonome convenable. On s'assure ainsi que *toutes les solutions de l'équation de Jacobi pour les écarts normaux sur une variété de courbure négative croissent, lorsqu'on se déplace sur une géodésique, au moins aussi vite que l'exponentielle du chemin parcouru dont l'exposant est égal à la racine carrée du module de la courbure sectionnelle qui minimise ce module.*

En réalité la plupart des solutions croissent encore plus vite, mais nous ne sommes plus en mesure d'affirmer que l'exposant de croissance de la plupart des solutions est défini par la direction de la plus grande courbure négative en module.

En résumé donc sur une variété de courbure négative les géodésiques sont exponentiellement instables. Pour évaluer cette instabilité introduisons le *chemin caractéristique* s comme le chemin moyen sur lequel les petites erreurs commises sur les conditions initiales croissent de e fois.

Plus exactement, le chemin caractéristique s peut être défini comme l'inverse de l'exposant λ qui caractérise l'accroissement des solutions de l'équation de Jacobi pour les écarts normaux par rapport à la géodésique parcourue à la vitesse unité:

$$\lambda = \lim_{T\to\infty} \frac{1}{T} \max_{|t|<T} \max_{|\xi(0)|=1} \ln |\xi(t)|,$$

$$s = 1/\lambda.$$

D'une façon générale, l'exposant λ et le chemin s dépendent de la géodésique initiale.

Si les courbures sectionnelles de notre variété sont $\leqslant -b^2$, *alors le chemin caractéristique n'est pas supérieur à* b^{-1}.

Donc plus la courbure négative est élevée en valeur absolue, plus est petit le chemin caractéristique s sur lequel l'instabilité des géodésiques entraîne un accroissement des erreurs de e fois. L'accroissement des erreurs étant exponentiel, il est pratiquement impossible de prédire la marche d'une géodésique sur une variété de courbure négative.

Supposons que la courbure est négative et $\leqslant -4$ m^{-2}. Le chemin caractéristique n'est pas supérieur à un demi-mètre, i.e. sur une portion de géodésique longue de cinq mètres l'erreur croît environ de e^{10}, soit de 10^4 fois. Donc une erreur d'un dixième de millimètre sur les conditions initiales se traduira par un écart de l'ordre du mètre au bout de la géodésique.

J. Flots géodésiques sur des variétés compactes de courbure négative. Soit M une variété riemannienne compacte de courbure sectionnelle négative en chaque point (ces variétés existent). Considérons le mouvement par inertie d'un point matériel de masse unité sur M en l'absence de forces extérieures. La fonction de Lagrange de ce système est égale à l'énergie cinétique, à l'énergie totale et est intégrale première des équations de mouvement.

Si la variété M est de dimension n, alors la variété de niveau d'énergie est de dimension $(2n - 1)$. C'est une sous-variété du fibré tangent à la variété M. Fixons par exemple la valeur de la constante d'énergie 1/2 (ce qui correspond à une vitesse initiale unité). La

vitesse du point restera toujours égale à l'unité et notre variété de niveau sera l'espace du fibré

$$T_1M \subset TM$$

constitué des sphères unités des espaces tangents à M en chacun de ses points.

Donc un point de la variété T_1M est un vecteur de longueur unité appliqué en un point de M. En vertu du principe de Maupertuis-Jacobi le mouvement d'un point matériel avec des conditions initiales fixées sur T_1M peut être décrit de la manière suivante: le point se déplace avec une vitesse unité sur la géodésique définie par le vecteur indiqué.

En vertu de la loi de conservation de l'énergie la variété T_1M est invariante dans l'espace des phases de notre système.

Donc notre flot définit un groupe à un paramètre de difféomorphismes de la variété T_1M de dimension $(2n - 1)$. Ce groupe est appelé *flot géodésique* sur M. Le flot géodésique peut être interprété comme suit: c'est une transformation qui pendant le temps t envoie un vecteur unité $\xi \in T_1M$ appliqué en un point x dans le vecteur vitesse unité de la géodésique issue du point x dans la direction de ξ et appliqué en un point de la géodésique se trouvant à une distance t du point x. On remarquera que l'élément de volume se définit naturellement sur T_1M et qu'il est invariant par le flot géodésique en vertu du théorème de Liouville.

La négativité de la courbure de la variété M a jusqu'ici été sans importance. Mais si l'on étudie les trajectoires du flot géodésique décrit, on s'apercevra qu'elle influence fortement le comportement de ces trajectoires (ceci est lié à l'instabilité exponentielle des géodésiques sur M).

Citons quelques propriétés des flots géodésiques sur des variétés de courbure négative (pour plus de détail voir l'ouvrage de D. Anossov cité à la page 300).

1. Toutes les trajectoires (à l'exception d'un ensemble de mesure nulle) sont partout denses dans une variété de niveau d'énergie.

2. Distribution uniforme: le temps de séjour de toutes les trajectoires (à l'exception d'un ensemble de mesure nulle) dans un domaine quelconque de l'espace des phases T_1M est proportionnel au volume de ce domaine.

3. Le flot g^t est « mixing », i.e. si A et B sont deux domaines, alors

$$\lim_{t \to \infty} \text{mes}\,[(g^t A) \cap B] = \text{mes}\, A \,\text{mes}\, B$$

(où « mes » désigne le volume normé par la condition que la mesure de l'espace tout entier est égale à l'unité).

Ces propriétés des trajectoires entraînent des propositons analogues pour les géodésiques sur la variété même. Les physiciens quali-

fient ces propriétés de « stochasticité » : une trajectoire asymptotique pour les grands t se conduit comme si le point était aléatoire. Par exemple, la propriété d'être « mixing » signifie que la probabilité de se retrouver en B à la sortie de A au bout d'un temps t grand est proportionnelle au volume de B, etc.

Ainsi, l'*instabilité exponentielle des géodésiques sur une variété de courbure négative entraîne la stochasticité du flot géodésique correspondant.*

K. Autres applications de l'instabilité exponentielle. L'instabilité exponentielle des géodésiques de courbure négative a été étudiée par de nombreux auteurs à commencer par Hadamard (et par Lobatchevski dans le cas d'une courbure constante) et en particulier par E. Hopf. Les années 60 ont été marquées par la découverte inattendue de l'étonnante stabilité des systèmes exponentiellement instables par rapport aux perturbations du système.

Considérons par exemple un champ de vecteurs définissant un flot géodésique sur une surface compacte de courbure négative. On sait que les courbes de phase de ce flot ont une allure très compliquée : chacune d'elles (à l'exception d'un ensemble de mesure nulle) est partout dense dans une variété de niveau d'énergie. Ce flot possède une infinité de trajectoires fermées et l'ensemble des points des trajectoires fermées est également partout dense dans une variété de niveau d'énergie de dimension trois.

Envisageons maintenant un champ de vecteurs proche. Il s'avère que le passage à ce champ ne modifie pratiquement pas ce tableau. Plus exactement, il existe un homéomorphisme proche de la transformation identique qui envoie les courbes de phase du flot non perturbé dans celles du flot perturbé.

Donc notre flot possède la même propriété de « grossièreté » ou de « stabilité structurelle » que, par exemple, un cycle limite, ou un foyer stable sur le plan. Remarquons que ni un centre sur le plan ni les hélices d'un tore ne jouissent de cette propriété : le type topologique des courbes de phase varie dans ces cas sous la moindre modification du champ de vecteurs.

L'éventualité de systèmes structurellement stables à mouvements compliqués dont chacun est exponentiellement instable en soi est à mettre au rang des plus importantes découvertes faites ces dernières années en théorie des équations différentielles (l'hypothèse de la stabilité structurelle des flots géodésiques sur des variétés de courbure négative a été avancée par S. Smale en 1961, la démonstration en a été faite par D. Anossov qui l'a publiée en 1967 ; J. Sinaï et D. Anossov ont obtenu des résultats fondamentaux sur la stochasticité de ces flots également dans les années soixante).

Jadis on supposait que dans les systèmes d'équations différentielles génériques ne pouvaient exister que des régimes limites

stables simples : des positions d'équilibre et des cycles. Si le système était plus compliqué (conservatif par exemple), on admettait que sous l'effet d'une faible modification des équations (par exemple si l'on tenait compte des petites perturbations non conservatives) les mouvements compliqués se « décomposaient » en mouvements simples. Maintenant nous savons qu'il en va autrement et que dans l'espace fonctionnel des champs de vecteurs il existe des domaines composés de champs où les courbes de phase sont plus complexes.

Les conclusions qui en découlent couvrent un grand nombre de phénomènes dans lesquels les objets déterministes ont un comportement « stochastique ».

Plus exactement, imaginons-nous que dans l'espace des phases d'un système (non conservatif) il existe une variété invariante attractive (ou un ensemble) sur laquelle les courbes de phase sont exponentiellement instables. Nous savons maintenant que les systèmes qui possèdent cette propriété ne sont pas une exception : les propriétés indiquées plus haut se conservent lorsque le système subit de petites variations. Par quoi sera surpris un expérimentateur observant le mouvement d'un tel système ?

Le rapprochement des courbes de phase de la variété attractive sera perçu comme l'établissement d'un régime limite. Le mouvement ultérieur du point représentatif au voisinage de la variété attractive provoquera dans le régime limite des variations de « phases » chaotiques, échappant à toute prévision qui seront regardées comme une « stochasticité» ou une « turbulence ».

Malheureusement à ce jour aucune analyse tant soit peu convaincante des exemples physiques de cette nature n'a été réalisée sous l'angle indiqué. L'exemple qui s'impose immédiatement c'est l'instabilité hydrodynamique d'un fluide visqueux décrit par l'équation de Navier-Stokes.

L'espace des phases de ce problème est de dimension infinie (c'est l'espace des champs de vecteurs de divergence zéro dans le domaine de l'écoulement), mais cette infinité de la dimension n'est visiblement pas un sérieux obstacle pour la bonne raison que la viscosité amortit les harmoniques supérieurs (les petits tourbillons) d'autant plus vite que l'ordre de l'harmonique est élevé. Par suite les courbes de phase de l'espace de dimension infinie se rapprochent de la variété de dimension finie (ou de l'ensemble) à laquelle appartiennent les régimes limites.

Lorsque la viscosité est grande, il existe une position d'équilibre attractive stable dans l'espace des phases (« écoulement stationnaire stable »).

Lorsque la viscosité diminue, cette position perd de sa stabilité ; et il peut apparaître, par exemple, un cycle limite stable dans l'espace des phases (« écoulement périodique ») ou une position d'équilibre stable d'un nouveau type (« écoulement stationnaire

secondaire »)*). Puis au fur et à mesure que la viscosité baisse, entrent en jeu un plus grand nombre d'harmoniques et les régimes limites peuvent acquérir une plus grande dimension.

Lorsque la viscosité est petite, il semble logique que l'on ait des régimes limites avec des trajectoires exponentiellement instables. Malheureusement les calculs nécessaires n'ont pu encore être effectués en raison des performances limitées des ordinateurs. Cependant on peut tirer une conclusion générale sans recourir au calcul, à savoir que les phénomènes de type turbulence n'impliquent pour se manifester ni la non-existence ni la non-unicité des solutions : il suffit de l'instabilité exponentielle que l'on rencontre même dans les systèmes déterministes à nombre fini de degrés de liberté.

Comme autre exemple d'application de l'instabilité exponentielle citons la démonstration exhibée par J. Sinaï de l'« hypothèse ergodique » de Boltzmann pour un système de billes solides. Cette hypothèse stipule que le flot correspondant au mouvement de billes parfaitement élastiques identiques dans un tiroir à parois élastiques est ergodique sur des ensembles connexes de niveau d'énergie. (L'ergodicité signifie que toutes les courbes de phase, à l'exception d'un ensemble de mesure nulle, restent dans chaque région mesurable d'un ensemble de niveau pendant un temps proportionnel à la mesure de cette région.)

L'hypothèse de Boltzmann permet de remplacer les moyennes temporelles par les moyennes spatiales et pendant longtemps a été considérée comme nécessaire à la justification de la mécanique statistique. En réalité, lorsqu'on passe à la limite en statistique (le nombre de particules tend vers l'infini), l'hypothèse de Boltzmann (où il est question de limite lorsque le temps tend vers l'infini) est inutile. Cependant cette hypothèse est à l'origine de toute l'analyse des propriétés stochastiques des systèmes dynamiques (la théorie ergodique) et sa démonstration mesure en quelque sorte la maturité de cette théorie.

L'instabilité exponentielle des trajectoires résulte dans le problème de Boltzmann des chocs entre les billes et peut être interprétée de la manière suivante.

Considérons par souci de simplicité un système composé de deux particules sur le plan et remplaçons le tiroir par un tore plan $\{(x, y) \bmod 1\}$. Nous pouvons alors considérer qu'une particule est fixe (en se servant de la conservation de l'impulsion) et assimiler l'autre à un point.

Nous sommes donc conduits à étudier le problème type du mouvement d'un point dans un billard torique à paroi centrale circulaire

*) Pour de plus amples détails sur la perte de stabilité consulter les *Cours sur les bifurcations et les familles verselles* (en russe). « Ouspekhi matematitcheskikh naouk », 27, nº 5 (1972), 119—184.

sur laquelle le point se réfléchit suivant la loi : l'angle d'incidence est égal à l'angle de réflexion (fig. 235).

Pour étudier ce système considérons un autre billard analogue borné extérieurement par une courbe convexe plane (par exemple le mouvement d'un point à l'intérieur d'une ellipse). Sur ce billard le mouvement peut être considéré comme le cas limite d'un flot

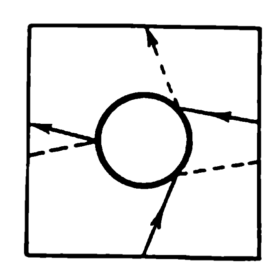


Fig. 235. Billard torique à paroi circulaire.

géodésique sur la surface d'un ellipsoïde. Le passage à la limite consiste à faire tendre le demi petit axe de l'ellipsoïde vers zéro. Par suite les géodésiques de l'ellipsoïde se transforment en trajectoires de la boule du billard sur l'ellipse. Ceci étant, on constate qu'il est plus logique de considérer l'ellipse comme étant à deux faces et qu'après chaque réflexion la géodésique passe d'une face sur l'autre.

Revenons maintenant au billard torique. Les mouvements sur ce billard peuvent être considérés comme un cas limite du flot géodésique sur une surface différentiable. On obtient cette surface si l'on considère un tore avec un trou comme une surface à deux faces à laquelle on confère une certaine épaisseur en arrondissant l'arête aiguë. En définitive on obtient une surface munie de la topologie d'une sphère à deux anses.

Si en « gonflant » une ellipse on obtient un ellipsoïde, i.e. une surface de courbure positive, en gonflant un tore comportant un trou on obtient une surface de courbure négative (dans les deux cas la courbure est concentrée à proximité du bord, mais le gonflage peut s'effectuer sans que le signe de la courbure change).

Donc dans le billard torique le mouvement peut être considéré comme le cas limite du mouvement sur une géodésique d'une surface de courbure négative.

Pour démontrer l'hypothèse de Boltzmann (dans le cas simple considéré) il suffit de vérifier que l'analyse des propriétés stochastiques des flots géodésiques sur les surfaces de courbure négative vaut dans le cas limite indiqué.

La démonstration détaillée est très compliquée ; seule a été publiée la partie qui concerne un système composé de deux particules (J. Sinaï, *Systèmes dynamiques à réflexions élastiques* (en russe). « Ouspekhi matematitcheskikh naouk », 25, n° 2 (1970), 141-192).

APPENDICE 2

GÉODÉSIQUES DES MÉTRIQUES INVARIANTES À GAUCHE SUR DES GROUPES DE LIE ET HYDRODYNAMIQUE DU FLUIDE PARFAIT

Le mouvement eulérien d'un solide peut être décrit comme un mouvement sur une géodésique sur le groupe des rotations d'un espace euclidien de dimension trois muni d'une métrique riemannienne invariante à gauche. Une partie importante de la théorie d'Euler n'est liée qu'à cette invariance et pour cette raison peut être transposée au cas d'autres groupes.

Parmi les exemples traités par cette théorie généralisée d'Euler distinguons le mouvement d'un solide dans un espace multidimensionnel et surtout l'hydrodynamique du fluide parfait (i.e. incompressible et non visqueux). Dans le dernier cas le rôle de groupe est assumé par un groupe de difféomorphismes d'un domaine d'écoulement, respectant l'élément de volume. Dans cet exemple le principe de moindre action signifie que le mouvement du liquide est décrit par une géodésique dans une métrique définie par l'énergie cinétique (si l'on veut, on peut adopter ce principe pour définition mathématique du fluide parfait). On s'assure sans peine que la métrique indiquée est invariante (à droite).

Avant de généraliser au cas infini les résultats obtenus pour les groupes de Lie de dimension finie il faut prendre quelques précautions. En effet les théorèmes d'existence et d'unicité des solutions des équations de mouvement n'ont pas encore été démontrés en hydrodynamique tridimensionnelle.

Néanmoins il est intéressant de voir ce que donne la généralisation formelle des propriétés des géodésiques sur des groupes de Lie de dimension finie. Les conclusions revêtent un caractère de propositions a priori (identités, inégalités, etc.) qui devraient être vraies moyennant une interprétation raisonnable des solutions. Parfois les conclusions ici obtenues sont susceptibles d'être justifiées rigoureusement et immédiatement sans passer par l'analyse du cas de dimension infinie.

Par exemple, les équations d'Euler du mouvement d'un solide ont pour analogue en hydrodynamique les équations d'Euler du mouvement du fluide parfait. Le théorème d'Euler relatif à la sta-

bilité des rotations autour du grand et du petit axe de l'ellipsoïde d'inertie a pour homologue le théorème légèrement généralisé de Rayleigh relatif à la stabilité des écoulements dont le profil des vitesses ne présente pas de points d'inflexion, etc.

Des formules d'Euler l'on déduit sans peine l'expression de la courbure riemannienne d'un groupe muni d'une métrique à invariance unilatérale. En les appliquant à l'hydrodynamique on trouve la courbure d'un groupe de difféomorphismes respectant l'élément de volume.

Il est intéressant de noter que pour les sections suffisamment bonnes la courbure est finie et dans nombre de cas, négative.

La négativité de la courbure est à l'origine de l'instabilité exponentielle des géodésiques (voir Appendice 1). Dans le cas considéré les géodésiques sont les lignes de courant du fluide parfait. Donc le calcul de la courbure du groupe de difféomorphismes nous renseigne partiellement sur l'instabilité des écoulements du fluide parfait.

C'est précisément la courbure qui définit le chemin caractéristique sur lequel les erreurs sur les conditions initiales croissent de e fois. La négativité de la courbure interdit pratiquement toute prédiction de l'écoulement : sur un chemin de plusieurs fois plus long que le chemin caractéristique l'écart des conditions initiales par rapport aux conditions initiales théoriques croît de plusieurs centaines de fois.

Dans cet Appendice on expose brièvement les résultats des calculs liés aux géodésiques sur des groupes munis d'une métrique à invariance unilatérale. Les démonstrations et développements figurent dans les travaux suivants :

1. V. Arnold — Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Annales de l'Institut Fourier, XVI, n° 1 (1966), 319-361.

2. V. Arnold — Sur une majoration a priori de la théorie de la stabilité en hydrodynamique (en russe). « Izvestia Institoutov », Mathématiques, n° 5 (54), 3-5.

3. V. Arnold — Remarques sur le comportement des écoulements des fluides parfaits tridimensionnels pour une petite perturbation du champ de vitesses initial (en russe). PMM 36, n° 2 (1972), 255-262.

4. V. Arnold — Hamiltonicité des équations d'Euler de la dynamique du solide et des fluides parfaits (en russe). OMN XXIX, n° 3 (147) (1969), 225-226.

5. L. Diki — Remarques sur les systèmes hamiltoniens attachés au groupe des rotations (en russe). « Analyse fonctionnelle et ses applications » 6, n° 4 (1972).

6. D. G. Ebin, J. Marsden — Groups of diffeomorphismes and the motion of an incompressible fluid. « Annals of Math. » 92, n° 1 (1970), 102-163.

7. O. L a d y j e n s k a ï a — Sur la résolubilité dans le petit des problèmes non stationnaires pour les fluides parfaits incompressibles et les fluides visqueux et pour une viscosité infime (en russe). Problèmes aux limites de la physique mathématique, t. 5. Léningrad, « Naouka », 1971, 65-78.

8. A. M i t c h e n k o — Intégrales des flots géodésiques sur les groupes de Lie (en russe). « Analyse fonctionnelle et ses applications » 4, n° 3 (1970), 73-78.

9. A. O b o u k h o v — Sur les invariants intégraux dans les systèmes de type hydrodynamique (en russe). DAN 184, n° 2 (1969).

10. L. F a d d é e v — Sur la théorie des écoulements parallèles plans stationnaires du fluide parfait (en russe). Problèmes aux limites de la physique mathématique, t. 5. Léningrad, « Naouka », 1971, 164-172.

A. Notations. Représentations adjointe et coadjointe. Soient G un groupe réel de Lie, $\mathfrak{g}$ son algèbre de Lie, i.e. l'espace tangent à G en son unité et muni de l'opération de commutation $[,]$.

Le groupe de Lie opère sur lui-même par translations à gauche et à droite : pour tout élément $g \in G$ sont définis des difféomorphismes du groupe sur lui-même

$$L_g : G \to G, \quad L_g h = gh; \quad R_g : G \to G, \quad R_g h = hg.$$

Les applications induites des espaces tangents seront notées

$$L_{g*} : TG_h \to TG_{gh}; \quad R_{g*} : TG_h \to TG_{hg}$$

pour tout $h \in G$.

Le difféomorphisme $R_{g^{-1}} L_g$ est un automorphisme intérieur du groupe. Il laisse l'unité du groupe en place. Sa dérivée en l'unité est une application linéaire de l'algèbre $\mathfrak{g}$ (i.e. de l'espace tangent au groupe en son unité) dans lui-même. Cette application sera notée

$$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{Ad}_g = (R_{g^{-1}} L_g)_{*e},$$

on l'appelle représentation *adjointe* du groupe.

On vérifie aisément que Ad_g est un homomorphisme d'algèbre, i.e.

$$\mathrm{Ad}_g [\xi, \eta] = [\mathrm{Ad}_g \xi, \mathrm{Ad}_g \eta], \quad \xi, \eta \in \mathfrak{g}.$$

Il est également clair que $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \mathrm{Ad}_h$.

Considérons par ailleurs Ad comme l'application du groupe G dans l'espace des opérateurs linéaires sur l'algèbre $\mathfrak{g}$: $\mathrm{Ad}(g) = \mathrm{Ad}_g$.

L'application Ad est différentiable. Etudions sa dérivée en l'unité du groupe G. Cette dérivée est une application linéaire de l'algèbre $\mathfrak{g}$ dans l'espace des opérateurs linéaires sur cette algèbre. Désignons cette application par ad ; alors ad_ξ sera l'image par cette application d'un élément $\xi \in \mathfrak{g}$. Donc ad_ξ est un endomorphisme d'algèbre

et nous avons

$$\mathrm{ad} = \mathrm{Ad}_{*e} : \mathfrak{g} \to \mathrm{End}\,\mathfrak{g}, \qquad \mathrm{ad}_\xi = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{e^{t\xi}},$$

où $e^{t\xi}$ est un groupe à un paramètre de vecteur tangent ξ. La formule précédente nous donne facilement l'expression de ad dans les seuls termes de notre algèbre

$$\mathrm{ad}_\xi \eta = [\xi, \eta].$$

Considérons l'espace vectoriel $\mathfrak{g}^*$ dual de l'algèbre de Lie $\mathfrak{g}$. C'est l'espace des fonctions réelles linéaires sur l'algèbre de Lie $\mathfrak{g}$. En d'autres termes, $\mathfrak{g}^*$ est l'espace cotangent au groupe G en son unité, $\mathfrak{g}^* = T^*G_e$.

La valeur d'un élément ξ de l'espace cotangent à G en l'un quelconque des points g sur un élément η de l'espace tangent en ce même point sera désignée par des parenthèses:

$$(\xi, \eta) \in \mathbf{R}, \qquad \xi \in T^*G_g, \qquad \eta \in TG_g.$$

Les translations à gauche et à droite induisent dans les espaces cotangents des opérateurs duaux de L_{g*} et R_{g*}. Nous les noterons

$$L_g^*: T^*G_{gh} \to T^*G_h, \qquad R_g^*: T^*G_{hg} \to T^*G_h$$

quel que soit $h \in G$. Ces opérateurs sont définis par les identités

$$(L_g^*\xi, \eta) \equiv (\xi, L_{g*}\eta), \qquad (R_g^*\xi, \eta) \equiv (\xi, R_{g*}\eta).$$

L'opérateur dual de Ad_g applique l'espace cotangent au groupe G en son unité dans lui-même. On le note

$$\mathrm{Ad}_g^*: \mathfrak{g}^* \to \mathfrak{g}^*$$

et on le définit au moyen de l'identité

$$(\mathrm{Ad}_g^*\xi, \eta) \equiv (\xi, \mathrm{Ad}_g \eta).$$

Les opérateurs Ad_g^*, où g parcourt le groupe de Lie G, forment une représentation de ce groupe, i.e. est vraie la relation

$$\mathrm{Ad}_{gh}^* = \mathrm{Ad}_h^* \mathrm{Ad}_g^*.$$

Cette représentation s'appelle *représentation coadjointe* du groupe de Lie, elle joue un rôle important dans tout ce qui touche aux métriques invariantes (à gauche) sur le groupe.

Considérons la dérivée de l'opérateur Ad_g^* par rapport à g en l'unité du groupe. Cette dérivée est une application linéaire de l'algèbre $\mathfrak{g}$ de l'espace des opérateurs linéaires sur l'espace $\mathfrak{g}^*$ dual de cette algèbre. Cette application se note ad^*; on désigne par ad_ξ^* l'image par cette application d'un élément $\xi \in \mathfrak{g}$. Donc ad_ξ^* est un opérateur linéaire sur l'espace dual de cette algèbre

$$\mathrm{ad}_\xi^*: \mathfrak{g}^* \to \mathfrak{g}^*.$$

On voit immédiatement que l'opérateur ad_ξ^* est conjugué de ad_ξ:

$$(ad_\xi^* \eta, \zeta) \equiv (\eta, ad_\xi \zeta) \text{ quels que soient } \eta \in g^*, \zeta \in g.$$

Quelquefois il est commode de représenter l'effet de ad^* par des accolades

$$ad_\xi^* \eta = \{\xi, \eta\}, \text{ où } \xi \in g, \eta \in g^*.$$

Donc l'accolade désigne une fonction bilinéaire de $g \times g^*$ dans g^* liée au commutateur par l'identité

$$(\{\xi, \eta\}, \zeta) = (\eta, [\xi, \zeta]).$$

Considérons les orbites de la représentation coadjointe du groupe G dans l'espace g^*. Sur chacune de ces orbites est définie une structure symplectique naturelle (appelée forme de Kirillov qui le premier s'en servit pour étudier les représentations des groupes nilpotents de Lie). Donc les orbites de la représentation coadjointe sont toujours de dimension paire. Remarquons par ailleurs que nous obtenons plusieurs exemples de variétés symplectiques en considérant divers groupes de Lie et l'ensemble de toutes les orbites.

On définit de la manière suivante une structure symplectique sur les orbites de la représentation coadjointe. Soient x un point de l'espace g^*, ξ le vecteur tangent en ce point à son orbite. Comme g^* est un espace vectoriel, nous pouvons considérer que le vecteur $\xi \in g^*$, bien qu'en réalité il appartienne à l'espace tangent à g^* en $x \in g^*$.

Le vecteur ξ peut être (de plusieurs manières) représenté sous forme du vecteur vitesse du mouvement du point x dans la représentation coadjointe d'un groupe à un paramètre e^{at} de vecteur vitesse $a \in g$. En d'autres termes, tout vecteur ξ tangent à l'orbite du point x dans la représentation coadjointe du groupe s'exprime en fonction d'un vecteur adéquat a de l'algèbre g au moyen de la formule

$$\xi = \{a, x\}, \quad a \in g, \quad x \in g^*.$$

Maintenant nous sommes à même de calculer la valeur d'une 2-forme symplectique Ω sur un couple de vecteurs ξ_1, ξ_2 tangents à l'orbite de x. Plus exactement, il faut exprimer ξ_1 et ξ_2 en fonction d'éléments quelconques a_1 et a_2 de l'algèbre g au moyen de la formule précédente, puis avec ces deux éléments et un troisième élément de l'algèbre duale g^* former le scalaire

$$\Omega(\xi_1, \xi_2) = (x, [a_1, a_2]), \quad x \in g^*, \quad a_i \in g.$$

On vérifie immédiatement que : 1) la forme bilinéaire Ω est correctement définie, i.e. sa valeur ne dépend pas du choix de a_i ; 2) que la forme Ω est antisymétrique et partant définit une 2-forme différentielle Ω sur l'orbite ; 3) que la forme Ω est non dégénérée et fermée (les démonstrations figurent dans l'appendice 5).

Donc la forme Ω est une structure symplectique sur toute orbite de la représentation coadjointe.

B. Métriques invariantes à gauche. Une métrique riemannienne sur un groupe de Lie G est par définition *invariante à gauche* si elle est invariante par toutes les translations à gauche L_g, i.e. si la dérivée d'une translation à gauche envoie chaque vecteur dans un vecteur de même longueur.

Il suffit de définir une métrique riemannienne invariante à gauche en un point, par exemple en l'unité : dans les autres points on peut alors la définir par des translations à gauche. Donc sur un groupe il y a autant de métriques riemanniennes invariantes à gauche que de structures euclidiennes sur l'algèbre $\mathfrak{g}$.

Une structure euclidienne sur l'algèbre $\mathfrak{g}$ est définie par un opérateur symétrique défini positif de l'algèbre $\mathfrak{g}$ dans l'espace dual $\mathfrak{g}^*$. Soit donc $A : \mathfrak{g} \to \mathfrak{g}^*$ un opérateur symétrique linéaire et défini positif :

$$(A\xi, \eta) = (A\eta, \xi) \text{ quels que soient } \xi, \eta \in \mathfrak{g}.$$

(La positivité de A n'est pas essentielle, mais dans les applications à la mécanique la forme quadratique $(A\xi, \xi)$ est définie positive.)

Définissons l'opérateur symétrique $A_g : TG_g \to T^*G_g$ au moyen de la translation à gauche

$$A_g\xi = L^*_{g^{-1}} A L_{g^{-1}*}\xi.$$

Nous obtenons donc le diagramme commutatif suivant d'opérateurs linéaires :

$$\begin{array}{ccccc}
\mathfrak{g}^* & & \xrightarrow{\mathrm{Ad}\,g} & & \\
{\scriptstyle A}\downarrow & \xleftarrow[L_{*g^{-1}}]{} & TG_g & \xrightarrow[R_{*g^{-1}}]{} & \mathfrak{g} \\
 & & \downarrow{\scriptstyle A_g} & & \\
\mathfrak{g}^* & \xleftarrow[L^*_g]{} & T^*G_g & \xrightarrow[R^*_g]{} & \mathfrak{g}^* \\
 & & \xleftarrow{\mathrm{Ad}^*_g} & &
\end{array}$$

Le produit scalaire défini par l'opérateur A_g sera noté par des chevrons

$$\langle \xi, \eta \rangle_g = (A_g\xi, \eta) = (A_g\eta, \xi) = \langle \eta, \xi \rangle_g.$$

Ce produit scalaire définit sur le groupe G une métrique riemannienne invariante par les translations à gauche.

Désignons le produit scalaire sur l'algèbre $\mathfrak{g}$ simplement par $\langle , \rangle$. Définissons l'opération $B : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ au moyen de l'identité

$$\langle [a, b], c \rangle \equiv \langle B(c, a), b \rangle \text{ quel que soit } b \in \mathfrak{g}.$$

Il est clair que l'opération B est bilinéaire et, le premier argument étant fixé, antisymétrique en le second:

$$\langle B(c, a), b\rangle + \langle B(c, b), a\rangle = 0.$$

C. Exemple. Soit $G = SO(3)$ le groupe des rotations de l'espace euclidien ordinaire, i.e. l'espace de configuration d'un solide ayant un point fixe. Le mouvement de ce solide est alors décrit par une courbe $g = g(t)$ sur le groupe G. L'algèbre de Lie du groupe G est l'espace tridimensionnel des vitesses angulaires de toutes les rotations. Le commutateur de cette algèbre est le produit vectoriel usuel.

La vitesse de rotation $\dot{g}$ du solide est un vecteur tangent au groupe G au point g. Pour obtenir la vitesse angulaire il faut transporter ce vecteur dans l'espace tangent au groupe en l'unité, i.e. dans l'algèbre $\mathfrak{g}$. On le fait de deux façons: par une translation à gauche ou par une translation à droite. On obtient ainsi deux vecteurs de l'algèbre $\mathfrak{g}$:

$$\omega_s = L_{g^{-1}*}\dot{g} \in \mathfrak{g}, \quad \omega_e = R_{g^{-1}*}\dot{g} \in \mathfrak{g}.$$

Ces deux vecteurs ne sont autres que la « vitesse angulaire par rapport au solide » et la « vitesse angulaire par rapport à l'espace ».

En effet, à tout élément g du groupe G correspond une position du solide qui s'obtient à partir d'un état initial (choisi arbitrairement et correspondant à l'unité du groupe) par le mouvement de g. Soit ω un élément de l'algèbre $\mathfrak{g}$.

Désignons par $e^{\omega t}$ le groupe à un paramètre de rotations de vitesse angulaire ω; ω est tangent en l'unité à ce groupe. Soit maintenant le déplacement

$$e^{\omega\tau}g, \text{ où } g = g(t) \in G, \ \omega \in \mathfrak{g}, \ \tau \ll 1,$$

déduit du déplacement de g par une rotation de vitesse angulaire ω pendant le temps τ petit. Si le vecteur $\dot{g}$ est confondu avec le vecteur

$$\left.\frac{d}{d\tau}\right|_{\tau=0} e^{\omega\tau}g,$$

on dit que ω est la *vitesse angulaire par rapport à l'espace* et on la note ω_e. Donc ω_e se déduit de $\dot{g}$ par une translation à droite. On démontre de façon analogue que la vitesse angulaire par rapport au solide est une translation à gauche du vecteur $\dot{g}$ dans l'algèbre $\mathfrak{g}$.

L'espace $\mathfrak{g}^*$ dual de l'algèbre $\mathfrak{g}$ est ici l'espace des moments cinétiques.

L'énergie cinétique du solide est définie par le vecteur vitesse angulaire par rapport au solide et ne dépend pas de la disposition du solide dans l'espace. Donc l'énergie cinétique définit une métrique riemannienne *invariante à gauche* sur le groupe G. L'opérateur symétrique défini positif $A_g: TG_g \to T^*G_g$ qui définit cette métrique s'appelle *opérateur* (ou tenseur) *d'inertie*: il est lié à l'énergie

cinétique par la formule

$$T=\frac{1}{2}\langle \dot{g}, \dot{g}\rangle_g=\frac{1}{2}\langle \omega_s, \omega_s\rangle=\frac{1}{2}(A\omega_s, \omega_s)=\frac{1}{2}(A_g\dot{g}, \dot{g}),$$

où $A: \mathfrak{g} \to \mathfrak{g}^*$ est la valeur de A_g pour $g = e$.

L'image du vecteur $\dot{g}$ par l'opérateur d'inertie A_g s'appelle *moment cinétique* et se note $M = A_g\dot{g}$.

Le vecteur M appartient à l'espace cotangent au groupe G au point g et on peut le transporter dans l'espace cotangent au groupe en l'unité soit par des translations à gauche, soit par des translations à droite. Il s'introduit ainsi deux vecteurs

$$M_s = L_g^*M \in \mathfrak{g}^*,$$
$$M_e = R_g^*M \in \mathfrak{g}^*.$$

Ces deux vecteurs de $\mathfrak{g}^*$ ne sont autres que le moment cinétique M_s par rapport au solide et le moment cinétique M_e par rapport à l'espace. Ce dernier se déduit facilement de l'expression qui donne l'énergie cinétique en fonction du moment et de la vitesse angulaire

$$T=\frac{1}{2}(M_s, \omega_s)=\frac{1}{2}(M, \dot{g}).$$

En vertu du principe de moindre action, le mouvement d'inertie du solide (en l'absence de forces extérieures) est une géodésique sur le groupe des rotations muni de la métrique invariante à gauche indiquée plus haut.

Soient maintenant à considérer les géodésiques d'une métrique riemannienne invariante à gauche sur un groupe quelconque de Lie comme les mouvements d'un « solide généralisé » d'espace de configuration G. Un tel « solide de groupe G » est défini par son énergie cinétique, i.e. une forme quadratique définie positive sur l'algèbre de Lie de G. Plus exactement, nous assimilerons les géodésiques de la métrique invariante définie sur le groupe G par la forme quadratique $\langle \omega, \omega \rangle$ sur l'algèbre de G aux mouvements d'un solide de groupe G et d'énergie cinétique $\langle \omega, \omega \rangle/2$.

A chaque mouvement $t \mapsto g(t)$ du solide généralisé nous pouvons faire correspondre les quatre courbes suivantes:

$$t \mapsto \omega_s(t) \in \mathfrak{g}, \quad t \mapsto \omega_e(t) \in \mathfrak{g},$$
$$t \mapsto M_s(t) \in \mathfrak{g}^*, \quad t \mapsto M_e(t) \in \mathfrak{g}^*,$$

appelées mouvements du vecteur vitesse angulaire et du vecteur moment par rapport au solide et par rapport à l'espace.

Les équations différentielles auxquelles obéissent ces courbes ont été établies par Euler pour un solide ordinaire. Cependant elles sont valables dans le cas le plus général d'un groupe arbitraire G: nous les appellerons équations d'Euler pour le solide généralisé.

R e m a r q u e. La théorie ordinaire du solide identifie six espaces distincts de dimension trois: $\mathbf{R}^3$, $\mathbf{R}^{3*}$, $\mathfrak{g}$, $\mathfrak{g}^*$, TG_g, T^*G_g. La dimension de l'espace $\mathbf{R}^3$ dans lequel se déplace le solide et celle de l'algèbre de Lie de son groupe de mouvement sont égales par pure coïncidence pour $n = 3$; dans le cas général la dimension de $\mathfrak{g}$ est $n(n-1)/2$.

L'identification de l'algèbre de Lie $\mathfrak{g}$ à son espace dual $\mathfrak{g}^*$ admet une justification bien plus profonde. En effet, sur le groupe des rotations est définie une (en une seule à un facteur près) métrique riemannienne invariante bilatérale. Cette métrique définit une fois pour toutes un isomorphisme entre les espaces vectoriels $\mathfrak{g}$ et $\mathfrak{g}^*$ (ainsi qu'entre TG_g et T^*G_g). Elle permet par conséquent de considérer que les vecteurs vitesse angulaire et moment appartiennent à un même espace euclidien. L'identification transforme l'opération $\{,\}$ en commutateur d'algèbre pris avec le signe moins.

Tout groupe compact de Lie est muni d'une métrique invariante bilatérale, donc en étudiant les mouvements des solides à groupes compacts on peut identifier les espaces des vitesses angulaires et des moments. Cependant nous ne ferons pas cette identification, l'essentiel pour nous étant les applications à un groupe non compact (et même de dimension infinie) de difféomorphismes.

D. Equations d'Euler. Les résultats obtenus par Euler (dans le cas particulier $G = SO(3)$) peuvent être énoncés sous forme de théorèmes relatifs au mouvement des vecteurs vitesse angulaire et moment des solides généralisés de groupe G.

Théorème 1. *Le vecteur moment par rapport à l'espace est invariant par le mouvement:*

$$\frac{dM_e}{dt} = 0.$$

Théorème 2. *Le vecteur moment par rapport au solide obéit à l'équation d'Euler*

$$\frac{dM_s}{dt} = \{\omega_s, M_s\}.$$

La démonstration de ces théorèmes pour le solide généralisé est la même que pour le solide ordinaire.

R e m a r q u e 1. Le vecteur vitesse angulaire ω_s par rapport au solide s'exprime linéairement en fonction du vecteur moment M_s par rapport au solide moyennant l'opérateur inverse de l'opérateur d'inertie: $\omega_s = A^{-1}M_s$. Donc l'équation d'Euler peut être considérée comme une équation pour le seul vecteur moment par rapport au solide; son second membre est quadratique en M_s.

Nous pouvons exprimer ce résultat d'une autre façon. Considérons en effet le flot de notre solide. (La dimension de l'espace des

phases T^*G du solide est le double de la dimension n du groupe G ou de l'espace des moments $\mathfrak{g}^*$.) Donc ce flot qui est défini sur une variété de dimension $2n$ se factorise sur le flot défini par l'équation d'Euler dans l'espace vectoriel $\mathfrak{g}^*$.

On appelle factorisation d'un flot g^t défini sur une variété X sur un flot f^t défini sur une variété Y une application différentiable π de X sur Y associant aux mouvements g^t les mouvements f^t de telle sorte que soit commutatif le diagramme

$$\begin{array}{ccc} X & \xrightarrow{g^t} & X \\ \pi \downarrow & & \pi \downarrow \\ Y & \xrightarrow{f^t} & Y \end{array}, \quad \text{i.e.} \quad \pi g^t = f^t \pi.$$

Ici $X = T^*G$ est l'espace des phases du solide, $Y = \mathfrak{g}^*$ l'espace des moments cinétiques. L'application projective π: $T^*G \to \mathfrak{g}^*$ est définie par les translations à gauche ($\pi M = L_g^* M$ pour $M \in T^*G_g$). Par ailleurs g^t est le flot du solide considéré dans l'espace T^*G de dimension $2n$ et f^t le flot de l'équation d'Euler dans l'espace des moments $\mathfrak{g}^*$ de dimension n.

Autrement dit, le mouvement du vecteur moment par rapport au solide ne dépend que de la position initiale du vecteur moment par rapport au solide et non pas de la position du solide dans l'espace.

R e m a r q u e 2. On peut exprimer la loi de conservation du vecteur moment par rapport à l'espace en disant que toute composante de ce vecteur est invariante dans un système quelconque de coordonnées sur l'espace $\mathfrak{g}^*$. Nous obtenons ainsi l'ensemble des intégrales premières des équations du mouvement du solide. En particulier, à tout élément de l'algèbre de Lie $\mathfrak{g}$ correspond une fonction linéaire sur l'espace $\mathfrak{g}^*$ et par conséquent une intégrale première. Un calcul immédiat nous montre que les crochets de Poisson des intégrales premières qui sont définies par des fonctions sur $\mathfrak{g}^*$ seront également des fonctions sur $\mathfrak{g}^*$. Nous obtenons donc une extension (de dimension infinie) de l'algèbre de Lie $\mathfrak{g}$ qui est composée de toutes les fonctions sur $\mathfrak{g}^*$. L'algèbre de Lie $\mathfrak{g}$ est plongée dans cette extension en tant qu'algèbre de Lie de fonctions linéaires sur $\mathfrak{g}^*$. Il est évident que de toutes ces intégrales premières du flot défini sur l'espace des phases de dimension $2n$ seules n seront fonctionnellement indépendantes. Pour ces n intégrales indépendantes on peut par exemple prendre n fonctions linéaires sur $\mathfrak{g}^*$ engendrant une base dans $\mathfrak{g}$.

S'agissant des applications au cas infini, il serait intéressant de se débarrasser des coordonnées et de formuler intrinsèquement la proposition relative aux intégrales premières. Ce qu'on peut faire en énonçant le théorème 1 sous la forme suivante.

Théorème 3. *Les orbites de la représentation coadjointe d'un groupe de l'espace dual de l'algèbre* $\mathfrak{g}$ *sont des variétés invariantes par le flot défini par l'équation d'Euler sur cet espace.*

D é m o n s t r a t i o n. $M_s(t)$ se déduit de $M_e(t)$ par la représentation coadjointe, quant à $M_e(t)$ il est fixe, c.q.f.d.

E x e m p l e. Dans le cas du solide ordinaire les orbites de la représentation coadjointe du groupe de l'espace des moments sont les sphères $M_1^2 + M_2^2 + M_3^2 = \text{const}$. Le théorème 3 se transforme donc en la loi de conservation du carré du moment. Ce théorème dit que si le point initial de M_s est sur une orbite quelconque (i.e. dans notre cas sur une sphère $M^2 = \text{const}$), alors tous les points de sa trajectoire sont situés sur la même orbite sous l'effet de l'équation d'Euler.

Revenons maintenant au cas général d'un groupe arbitraire de Lie et souvenons-nous que les orbites de la représentation coadjointe possèdent une structure symplectique (voir point A). D'autre part l'énergie cinétique du solide peut être exprimée en fonction du moment par rapport au solide. En définitive nous obtenons la forme quadratique sur l'espace des moments

$$T = \frac{1}{2}(M_s, A^{-1}M_s).$$

Fixons une orbite quelconque V de la représentation coadjointe. Considérons l'énergie cinétique comme une fonction sur cette orbite :

$$H : V \to \mathbf{R}, \quad H(M_s) = \frac{1}{2}(M_s, A^{-1}M_s).$$

Théorème 4. *Sur chaque orbite V de la représentation coadjointe l'équation d'Euler est hamiltonienne de fonction de Hamilton H.*

D é m o n s t r a t i o n. Tout vecteur ξ tangent à V en un point x est de la forme $\xi = \{f, M\}$, où $f \in \mathfrak{g}$. En particulier le champ de vecteurs du second membre de l'équation d'Euler peut se mettre sous la forme $X = \{dT, M\}$ (la différentielle de la fonction T en $M \in \mathfrak{g}^*$ est considérée ici comme un vecteur de l'espace dual de $\mathfrak{g}^*$, i.e. comme un élément de l'algèbre de Lie $\mathfrak{g}$). Des définitions de la structure symplectique Ω et de l'opération $\{,\}$ (voir point A) il résulte que pour tout vecteur ξ tangent en M à V, l'on a

$$\Omega(\xi, X) = (M, [f, dT]) = (dT, \{f, M\}) = (dH, \xi),$$

c.q.f.d.

L'équation d'Euler peut être transférée de l'espace $\mathfrak{g}^*$ dans l'algèbre $\mathfrak{g}$ par inversion de l'opérateur d'inertie. On obtient ainsi la formulation suivante pour l'équation d'Euler dans les termes de l'opération B (page 323).

Théorème 5. *Le mouvement du vecteur vitesse angulaire par rapport au solide est défini par la position initiale de ce vecteur et ne dépend pas de celle du solide. Le vecteur vitesse angulaire par rapport au solide satisfait l'équation de second membre quadratique*

$$\dot{\omega}_s = B(\omega_s, \omega_s).$$

Cette équation est l'équation d'Euler pour la vitesse angulaire. On remarquera que les orbites de la représentation coadjointe se transforment par l'opérateur A^{-1}: $\mathfrak{g}^* \to \mathfrak{g}$ en variétés invariantes de l'équation d'Euler pour la vitesse angulaire: ces variétés possèdent une structure symplectique, etc. Cependant à la différence des orbites de $\mathfrak{g}^*$ ces variétés invariantes ne sont pas définies directement par le groupe de Lie G, mais elles dépendent aussi du choix du solide (i.e. de l'opérateur d'inertie).

La loi de conservation de l'énergie entraîne le

Théorème 6. *Les équations d'Euler (pour les moments et les vitesses angulaires) admettent une intégrale première quadratique égale à l'énergie cinétique*

$$T = \frac{1}{2}(M_s, A^{-1}M_s) = \frac{1}{2}(A\omega_s, \omega_s).$$

E. Rotations stationnaires et stabilité. On appelle *rotation stationnaire* d'un solide une rotation de vitesse angulaire constante (par rapport au solide et par rapport à l'espace: on voit d'ailleurs que l'une entraîne l'autre). De la théorie du solide ordinaire dans $\mathbf{R}^3$ nous savons que les rotations stationnaires sont des rotations autour des axes principaux de l'ellipsoïde d'inertie. Plus bas on formule la généralisation de ce théorème au cas d'un solide de groupe de Lie quelconque. On remarquera que les rotations stationnaires sont les géodésiques d'une métrique invariante qui sont des sous-groupes à un paramètre.

Remarquons par ailleurs que l'on peut définir les directions des axes principaux de l'ellipsoïde d'inertie en considérant les points stationnaires de l'énergie cinétique sur une sphère des vecteurs moment de longueur fixée.

Théorème 7. *Le moment cinétique (resp. la vitesse angulaire) d'une rotation stationnaire par rapport au solide est un point critique de l'énergie sur une orbite de la représentation coadjointe (resp. sur l'image d'une orbite par l'opérateur A^{-1}). Réciproquement, tout point critique de l'énergie sur une orbite définit une rotation stationnaire.*

On démontre ce théorème soit par un calcul immédiat, soit en se référant au théorème 4.

Remarquons que la partition de l'espace des moments en orbites de la représentation coadjointe dans le cas d'un groupe arbitraire n'est pas forcément aussi simple que dans le cas élémentaire d'un solide ordinaire où, rappelons-le, nous avons obtenu des sphères de centre O et le point O. En effet, les orbites peuvent être de dimension différente et la partition en orbites peut ne pas être un fibré au voisinage du point donné: cette singularité se présente déjà dans le cas tridimensionnel au point O.

Nous dirons d'un point M de l'espace des moments qu'il est *régulier* si la partition d'un voisinage de ce point en orbites est difféomorphe à la partition de l'espace euclidien en plans parallèles (en particulier, toutes les orbites proches du point M ont même dimension). S'agissant par exemple du groupe des rotations de l'espace ordinaire, tous les points de l'espace des moments sont réguliers à l'exception de l'origine des coordonnées.

Théorème 8. *Supposons qu'un point régulier M de l'espace des moments est point critique de l'énergie cinétique sur une orbite de la représentation coadjointe et que la différentielle seconde d^2H de l'énergie est une forme quadratique définie positive ou négative. Sous ces conditions M est une position d'équilibre stable (au sens de Liapounov) de l'équation d'Euler.*

La démonstration se base sur le fait que l'énergie présente un maximum ou un minimum liés par suite de la régularité sur les orbites voisines du point M.

Théorème 9. *La différentielle seconde de l'énergie cinétique restreinte à l'image d'une orbite de la représentation coadjointe de l'algèbre $\mathfrak{g}$ est définie en un point $\omega \in \mathfrak{g}$ par la formule*

$$2d^2H|_\omega(\xi) = \langle B(\omega, f), B(\omega, f)\rangle + \langle [f, \omega], B(\omega, f)\rangle,$$

où ξ est un vecteur tangent à l'image indiquée s'exprimant en fonction de f par la formule

$$\xi = B(\omega, f), \quad f \in \mathfrak{g}.$$

F. Courbure riemannienne d'un groupe à métrique invariante à gauche. Soit G un groupe de Lie muni d'une métrique invariante à gauche que nous supposons être définie par le produit scalaire $\langle , \rangle$. Remarquons que la courbure riemannienne du groupe G en l'un quelconque de ses points est définie par la courbure en l'unité (puisque les translations à gauche transforment isométriquement le groupe en lui-même). Donc il suffit de calculer la courbure pour des 2-plans contenus dans l'algèbre de Lie.

Théorème 10. *La courbure du groupe G dans la section définie par un couple de vecteurs orthonormés ξ, η de l'algèbre $\mathfrak{g}$ est donnée par la formule*

$$K_{\xi,\eta} = \langle \delta, \delta\rangle + 2\langle \alpha, \beta\rangle - 3\langle \alpha, \alpha\rangle - 4\langle B_\xi, B_\eta\rangle,$$

où $2\delta = B(\xi, \eta) + B(\eta, \xi)$, $2\beta = B(\xi, \eta) - B(\eta, \xi)$, $2\alpha = [\xi, \eta]$, $2B_\xi = B(\xi, \xi)$, $2B_\eta = B(\eta, \eta)$

et B l'opération définie au point B (page 323).

La démonstration se fait par un calcul direct fastidieux qui consiste à vérifier la formule pour la dérivée covariante

$$(\nabla_\xi \eta)_e = \frac{1}{2}([\xi, \eta] - B(\xi, \eta) - B(\eta, \xi)),$$

où dans le premier membre ξ, η sont des champs de vecteurs invariants à gauche et dans le second membre, leurs valeurs en l'unité.

R e m a r q u e 1. Dans le cas particulier d'une métrique invariante bilatérale la formule de la courbure prend une forme très simple

$$K_{\xi, \eta} = \frac{1}{4}\langle [\xi, \eta], [\xi, \eta] \rangle.$$

R e m a r q u e 2. La formule de la courbure est la même que le groupe soit muni d'une métrique riemannienne invariante à gauche ou à droite. En effet, une métrique invariante à droite sur un groupe est invariante à gauche sur ce groupe si on permute la loi de multiplication ($g_1 * g_2 = g_2 g_1$). En permutant l'opération de multiplication on change les signes du commutateur et de l'opération B en même temps. Or chaque terme de la formule de la courbure est un produit de deux opérations qui changent de signe. Donc la formule de la courbure est la même dans le cas d'une métrique invariante à droite.

Lorsque la métrique est invariante à droite, le second membre de l'équation d'Euler change de signe.

G. Applications à un groupe de difféomorphismes. Soit D un domaine borné dans une variété riemannienne. Considérons le groupe de difféomorphismes de D qui conserve l'élément de volume. Désignons ce groupe par SDiff D.

L'algèbre de Lie correspondant au groupe SDiff D est composée de tous les champs de vecteurs de divergence nulle sur D et tangents au bord (s'il n'est pas vide). Définissons le produit scalaire de deux éléments de cette algèbre de Lie (i.e. de deux champs de vecteurs) par

$$\langle v_1, v_2 \rangle = \int_D (v_1 \cdot v_2)\, dx,$$

où $(\cdot)$ est le produit scalaire définissant une métrique riemannienne sur D et dx est l'élément de volume riemannien.

Considérons maintenant l'écoulement d'un fluide homogène parfait (i.e. incompressible et non visqueux) dans le domaine D. Cet écoulement est décrit par une courbe $t \mapsto g_t$ sur le groupe SDiff D. D'une façon plus précise un difféomorphisme g_t est une application qui transporte toute particule de fluide de l'endroit où elle se trouvait à la date 0 à l'endroit où elle se trouvera à la date t.

Il s'avère que l'énergie cinétique du fluide en mouvement est une métrique riemannienne invariante à droite sur le groupe de difféomorphismes SDiff D.

En effet, supposons que pendant le temps t l'écoulement du fluide a réalisé le difféomorphisme g_t et que la vitesse soit donnée à cet instant par le champ v. Alors le difféomorphisme opéré par l'écoulement pendant le temps $t+\tau$ (où τ est petit) sera $e^{v\tau}g_t$ aux infiniment petits par rapport à τ près ($e^{v\tau}$ est ici un groupe à un paramètre de vecteur vitesse v, i.e. le flot de l'équation différentielle définie par le champ v).

Donc le champ de vitesses v s'obtient à partir du vecteur $\dot{g}$ tangent au groupe en x par une translation *à droite*. De là résulte l'invariance à droite de l'énergie cinétique qui, par définition, est égale à

$$T=\frac{1}{2}\langle v, v\rangle$$

(la densité du fluide est supposée égale à l'unité).

Le principe de moindre action (qui mathématiquement est la définition du fluide parfait) affirme que les écoulements du fluide parfait sont les géodésiques de la métrique invariante à droite décrite sur un groupe de difféomorphismes.

En toute rigueur, un groupe de difféomorphismes de dimension infinie n'est pas une variété. Donc la formulation exacte de la définition précédente a besoin d'être complétée: il faut choisir des espaces fonctionnels convenables, démontrer le théorème d'existence et d'unicité des solutions, etc. Jusqu'à présent on n'a pu le faire que dans le cas où la dimension du domaine d'écoulement D était égale à deux. Pour notre part nous procéderons comme si les difficultés liées à l'infinité de la dimension n'existaient pas. Donc les raisonnements ultérieurs revêtiront un caractère heuristique. Au demeurant de nombreux résultats peuvent être rigoureusement démontrés indépendamment de la théorie des variétés de dimension infinie.

Voyons maintenant comment s'écrivent les formules générales exhibées plus haut dans le cas $G=$ SDiff D où D est un domaine connexe de volume fini dans une variété riemannienne de dimension trois. Pour cela il nous faut avant tout expliciter l'opération bilinéaire $B\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ définie dans le point B par l'identité

$$\langle[a, b], c\rangle\equiv\langle B(c, a), b\rangle.$$

On vérifie sans peine que dans le cas tridimensionnel le champ de vecteurs $B(c, a)$ s'exprime en fonction des champs de vecteurs a et c de l'algèbre de Lie $\mathfrak{g}$ au moyen de la formule

$$B(c, a)=(\operatorname{rot} c)\wedge a+\operatorname{grad}\alpha,$$

où $\wedge$ désigne le produit vectoriel, α une fonction univoque dans D qui est définie de façon unique (à l'addition près d'un terme) par la condition $B\in\mathfrak{g}$ (i.e. par la condition $\operatorname{div} B=0$ et B tangent au bord de D).

Remarquons que l'opération B ne dépend pas du choix de l'orientation, puisque le produit vectoriel et le rotationnel sont multipliés tous deux par -1 lorsque l'orientation change.

H. Ecoulements stationnaires. L'équation d'Euler pour la « vitesse » angulaire dans le cas $G = \mathrm{SDiff}\, D$ est de la forme $\dot{v} = -B(v, v)$, puisque la métrique est invariante à droite. Dans le cas d'un groupe de difféomorphismes de l'espace elle prend donc la forme de l'« équation du mouvement de Bernoulli »

$$\frac{\partial v}{\partial t} = v \wedge \mathrm{rot}\, v + \mathrm{grad}\, \alpha, \quad \mathrm{div}\, v = 0.$$

L'équation d'Euler pour le moment s'écrit sous forme de l'« équation du rotationnel »

$$\frac{\partial\, \mathrm{rot}\, v}{\partial t} = [v, \mathrm{rot}\, v].$$

En particulier, le *rotationnel d'un écoulement stationnaire commute avec le champ de vitesses.*

Cette remarque permet aussitôt de classer topologiquement les écoulements stationnaires du fluide parfait dans l'espace.

Théorème 11. *Supposons que le domaine D est compact et borné par une surface analytique, et que le champ de vitesses est analytique et pas partout colinéaire à son rotationnel. Alors le domaine d'écoulement est partagé par un sous-ensemble analytique en un nombre fini de cellules dans chacune desquelles l'écoulement est canonique. Plus exactement, les cellules sont de deux types : feuilletées en tores invariants par l'écoulement et en surfaces difféomorphes à la couronne* $\mathbf{R} \times S^1$ *et invariantes par l'écoulement. En outre les lignes de courant soit sont fermées, soit sont partout denses sur chacun des tores, soit encore sont fermées dans chaque couronne.*

La démonstration de ce théorème repose sur l'étude des « surfaces de Bernoulli », i.e. des surfaces de niveau de la fonction α. La condition de stationnarité ($v \wedge \mathrm{rot}\, v = -\mathrm{grad}\, \alpha$) entraîne que les lignes de courant comme les lignes caractéristiques sont situées sur les surfaces de Bernoulli. Puisque les champs de vitesses et du rotationnel commutent, le groupe $\mathbf{R}^2$ opère sur une surface fermée de Bernoulli qui est forcément un tore (comparer avec la démonstration du théorème de Liouville au § 49). Par des raisonnements analogues et compte tenu de la condition imposée au bord de D on démontre que les surfaces non fermées de Bernoulli sont constituées de couronnes dans lesquelles sont fermées les lignes de courant.

Remarque. L'analyticité du champ de vitesses n'est pas très essentielle, mais il est important que les champs de vitesses et du rotationnel ne soient colinéaires dans aucun domaine. Les calculs

faits sur ordinateur par M. Hanon montrent que les lignes de courant pour un écoulement stationnaire défini par les formules

$$v_x = A \sin z + C \cos y, \quad v_y = B \sin x + A \cos z,$$
$$v_z = C \sin y + B \cos x$$

sur un tore de dimension trois ont un comportement plus compliqué que celui prévu par le théorème. Les formules sont choisies de telle sorte que les vecteurs v et rot v soient colinéaires. A en juger par les résultats du calcul, certaines lignes de courant sont partout denses dans des domaines de dimension trois.

I. Champs isorotationnels. L'hydrodynamique bidimensionnelle se distingue foncièrement de l'hydrodynamique tridimensionnelle par la géométrie des orbites des représentations coadjointes. En effet, dans le cas bidimensionnel les orbites sont dans un certain sens fermées et se conduisent à peu près comme une famille d'ensembles de niveau d'une fonction (plus exactement de quelques fonctions, voire, en réalité, d'un nombre infini de fonctions). Dans le cas tridimensionnel les orbites sont plus compliquées et en particulier infinies (et éventuellement denses). Les orbites de la représentation coadjointe d'un groupe de difféomorphismes d'une variété riemannienne de dimension trois peuvent être décrites de la manière suivante. Soient v_1 et v_2 deux champs de vecteurs des vitesses d'un fluide incompressible sur le domaine D. Nous dirons que les champs v_1 et v_2 sont *isorotationnels* s'il existe un difféomorphisme $g: D \to D$ envoyant tout contour fermé γ de D dans un autre contour fermé de D, respectant l'élément de volume et tel que la circulation du premier champ le long du contour de départ soit égale à la circulation du second le long du contour d'arrivée:

$$\oint_{\gamma} v_1 = \oint_{g\gamma} v_2.$$

On vérifie aisément que l'image d'une orbite de la représentation coadjointe dans l'algèbre $\mathfrak{g}$ (par l'inverse A^{-1} de l'opérateur d'inertie) n'est autre chose qu'un ensemble de champs isorotationnels au champ donné.

En particulier, le théorème 3 s'énonce maintenant sous la forme de la *loi de conservation de la circulation.*

Théorème 12. *La circulation d'un champ de vecteurs d'un fluide parfait le long d'un contour liquide fermé ne change pas lorsque ce contour est transporté par le liquide.*

Remarquons que si deux champs de vitesses d'un fluide parfait tridimensionnel sur D sont isorotationnels, alors le difféomorphisme

correspondant envoie le rotationnel du premier dans celui du second :

$$g_* \operatorname{rot} v_1 = \operatorname{rot} v_2.$$

Bien plus, l'isorotationnalité de deux champs peut être définie comme l'équivalence des champs de rotationnels *si le domaine d'écoulement est simplement connexe*. Donc le problème relatif aux orbites de la représentation coadjointe dans le cas tridimensionnel généralise le problème relatif à la classification des *champs de vecteurs* de divergence nulle aux difféomorphismes près conservant l'élément de volume. Ce problème est incommensurablement compliqué dans le cas tridimensionnel.

Traitons maintenant le cas bidimensionnel. Tout d'abord écrivons les principales formules dans une forme commode à l'étude de ce cas.

Supposons que le domaine d'écoulement D est orienté et de dimension deux. La métrique et l'orientation définissent sur D une structure symplectique; le champ de vecteurs vitesse possède une divergence nulle, donc il est hamiltonien et par conséquent est défini par une fonction de Hamilton (en général multivoque si le domaine D est multiplement connexe). La fonction de Hamilton du champ de vitesses s'appelle *fonction de courant* en hydrodynamique et se note ψ. Donc

$$v = I \operatorname{grad} \psi,$$

où I est une rotation de 90° « à droite ».

La fonction de courant du commutateur de deux champs est le jacobien (ou si l'on veut le crochet de Poisson du formalisme hamiltonien) des fonctions de courant des champs de départ

$$\psi_{[v_1, v_2]} = J(\psi_1, \psi_2).$$

Le champ de vecteurs $B(c, a)$ est défini dans le cas bidimensionnel par la formule

$$B = -(\Delta\psi_c) \operatorname{grad} \psi_a + \operatorname{grad} \alpha,$$

où ψ_a et ψ_c sont les fonctions de courant des champs a et c, $\Delta =$ $=$ div grad, le laplacien.

Dans le cas particulier du plan euclidien muni des coordonnées cartésiennes x et y, les formules de la fonction de courant, du commutateur et du laplacien prennent la forme particulièrement simple

$$v_x = \frac{\partial\psi}{\partial y}, \quad v_y = -\frac{\partial\psi}{\partial x},$$

$$\psi_{[v_1, v_2]} = \frac{\partial\psi_{v_1}}{\partial x}\frac{\partial\psi_{v_2}}{\partial y} - \frac{\partial\psi_{v_1}}{\partial y}\frac{\partial\psi_{v_2}}{\partial x},$$

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Le rotationnel d'un champ de vitesses de dimension deux est une fonction scalaire dont l'intégrale du produit par l'élément de surface orienté sur un domaine orienté quelconque σ de D est égale à la circulation du champ de vitesses le long du bord de σ:

$$\int_\sigma r\, dS = \oint_{\partial\sigma} v.$$

On exprime immédiatement le rotationnel en fonction de ψ:

$$r = -\Delta\psi.$$

Dans un domaine simplement connexe de dimension deux, l'isorotationnalité des champs v_1 et v_2 signifie tout simplement que les fonctions r_1 et r_2 (i.e. les rotationnels de ces champs) se correspondent par un difféomorphisme convenable respectant l'élément de surface.

Deux fonctions r_1 et r_2 possédant cette propriété sont en tout cas de même mesure, i.e.

$$\text{mes}\,\{x \in D:\ r_1(x) \leqslant c\} = \text{mes}\,\{x \in D:\ r_2(x) \leqslant c\},$$

quel que soit le nombre c. Donc l'appartenance de deux champs à l'image d'une même orbite de la représentation coadjointe entraîne l'égalité de toute une chaîne de fonctionnelles, par exemple les intégrales de toutes les puissances du rotationnel

$$\int_D r_1^k\, dS = \int_D r_2^k\, dS.$$

En particulier, les équations d'Euler de mouvement d'un fluide parfait bidimensionnel

$$\frac{\partial v}{\partial t} + v\nabla v = -\text{grad}\, p, \quad \text{div}\, v = 0,$$

admettent une infinité d'intégrales premières. Par exemple, l'intégrale d'une puissance quelconque du rotationnel du champ de vitesses

$$I_k = \iint_D \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)^k dx \wedge dy$$

est également intégrale première.

C'est précisément l'existence de ces intégrales premières (i.e. la structure relativement simple des orbites de la représentation coadjointe) qui a permis de démontrer les théorèmes d'existence, d'unicité, etc. dans l'hydrodynamique bidimensionnelle d'un fluide parfait (et même visqueux) ; et c'est précisément à cause de la géométrie compliquée des orbites de la représentation coadjointe dans le cas tridimensionnel (ou plutôt à cause de l'insuffisance d'information

sur ces orbites) qu'il est extrêmement difficile de justifier l'hydrodynamique tridimensionnelle.

J. Stabilité des écoulements stationnaires plans. Formulons maintenant les théorèmes généraux relatifs aux rotations stationnaires (théorèmes 7, 8 et 9) pour le cas d'un groupe de difféomorphismes. Nous obtenons les propositions suivantes :

1. *L'écoulement stationnaire d'un fluide parfait se distingue de tous les écoulements qui lui sont isorotationnels par le fait qu'il est un point d'extremum lié (ou un point critique) de l'énergie cinétique.*

2. *Si 1° le point critique mentionné est réellement un extremum, i.e. est un maximum ou un minimum local lié, 2° est réalisée une condition de régularité et 3° l'extremum est non dégénéré (la différentielle seconde est définie positive ou négative), alors cet écoulement stationnaire est stable (i.e. est une position d'équilibre stable au sens de Liapounov de l'équation d'Euler).*

3. *La formule de la différentielle seconde de l'énergie cinétique sur un espace tangent à la variété des champs isorotationnels au champ donné s'écrit sous la forme suivante dans le cas bidimensionnet : soil D un domaine du plan euclidien muni des coordonnées x, y et soit $\psi = \psi(x, y)$ la fonction de courant d'un écoulement stationnaire. Alors*

$$2d^2H = \iint_D (\delta v)^2 + \frac{\nabla\psi}{\nabla\Delta\psi}(\delta r)^2\, dx\, dy,$$

où δv est la variation du champ de vitesses (i.e. un vecteur de l'espace tangent indiqué plus haut) et $\delta r = \operatorname{rot} \delta v$.

Remarquons que, pour un écoulement stationnaire, les vecteurs gradient de la fonction de courant et son laplacien sont colinéaires. Donc le rapport $\nabla\psi/\nabla\Delta\psi$ a un sens. D'autre part au voisinage de tout point où le gradient du rotationnel n'est pas nul, la fonction de courant est une fonction de la fonction du rotationnel.

Ces propositions nous permettent donc de conclure qu'une condition suffisante de stabilité de l'écoulement stationnaire envisagé est que la forme quadratique d^2H soit définie positive ou négative. Cette conclusion ne découle pas formellement des théorèmes 7, 8 et 9, puisque l'application de nos formules au cas de dimension infinie doit être prouvée.

Heureusement il est possible de démontrer la conclusion finale sur la stabilité sans avoir à justifier les constructions intermédiaires. On peut donc démontrer rigoureusement les majorations a priori suivantes (qui expriment la stabilité de l'écoulement stationnaire par rapport aux petites perturbations du champ de vitesses initial).

Théorème 13. *Supposons que la fonction de courant $\psi = \psi(x, y)$ d'un écoulement stationnaire dans un domaine D est une fonction de la fonction du rotationnel (i.e. de la fonction $-\Delta\psi$) non seulement locale-*

ment mais globalement. Supposons que la dérivée de la fonction de courant par rapport à la fonction du rotationnel vérifie l'inégalité

$$c \leqslant \frac{\nabla\psi}{\nabla\Delta\psi} \leqslant C, \quad \text{où} \quad 0 < c \leqslant C < \infty.$$

Soit $\psi + \varphi(x, y, t)$ *la fonction de courant d'un autre écoulement, pas forcément stationnaire. Supposons qu'à la date initiale la circulation du champ de vitesses de l'écoulement perturbé (de fonction de courant* $\psi + \varphi$*) le long de chaque composante du bord du domaine* D *est égale à la circulation de l'écoulement initial (de fonction de courant* ψ*). Sous toutes ces conditions la perturbation de* $\varphi = \varphi(x, y, t)$ *à une date quelconque est majorée par la perturbation initiale de* $\varphi_0 = \varphi(x, y, 0)$ *au moyen de la formule*

$$\iint_D (\nabla\varphi)^2 + c(\Delta\varphi)^2\, dx\, dy \leqslant \iint_D (\nabla\varphi_0)^2 + C(\Delta\varphi_0)^2\, dx\, dy.$$

Si l'écoulement stationnaire vérifie l'inégalité

$$c \leqslant -\frac{\nabla\psi}{\nabla\Delta\psi} \leqslant C, \quad 0 < c < C < \infty,$$

alors la perturbation de φ *est majorée par celle de* φ_0 *au moyen de la formule*

$$\iint_D c(\Delta\varphi)^2 - (\nabla\varphi)^2\, dx\, dy \leqslant \iint_D C(\Delta\varphi_0)^2 - (\nabla\varphi_0)^2\, dx\, dy.$$

Ce théorème entraîne la stabilité de l'écoulement stationnaire sous réserve que soit définie positive la forme quadratique en $\nabla\varphi$

$$\iint_D (\nabla\varphi)^2 + \frac{\nabla\psi}{\nabla\Delta\psi}(\Delta\varphi)^2\, dx\, dy$$

(où φ est une fonction constante sur chaque composante du bord de D et dont est nul le flux du gradient à travers chaque composante du bord de D) ou que soit définie négative la forme quadratique

$$\iint_D (\nabla\varphi)^2 + \left(\max \frac{\nabla\psi}{\nabla\Delta\psi}\right)(\Delta\varphi)^2\, dx\, dy.$$

Exemple 1. Considérons un écoulement parallèle plan dans la bande $Y_1 \leqslant y \leqslant Y_2$ du plan (x, y) de profil des vitesses $v(y)$ (i.e. de champ de vitesses $(v(y), 0)$). Un tel écoulement est stationnaire quel que soit le profil des vitesses. Pour rendre le domaine d'écoulement compact imposons au champ de vitesses de tous les écoulements considérés d'être périodique, de période X en la coordonnée x.

La condition du théorème 13 est réalisée si le profil des vitesses ne présente pas de points d'inflexion (i.e. si $d^2v/dy^2 \neq 0$). Nous con-

cluons donc que les *écoulements parallèles plans d'un fluide parfait sont stables si le profil des vitesses ne présente pas de points d'inflexion.*

La proposition analogue dans le problème linéarisé s'appelle *théorème de Rayleigh.*

A souligner que dans le théorème 13 il est question non pas d'une stabilité « dans une approximation linéaire », mais d'une stabilité rigoureuse au sens de Liapounov (i.e. par rapport aux perturbations finies dans le problème non linéaire). La différence entre ces deux types de stabilité est essentielle dans le cas considéré, puisque notre problème est hamiltonien (voir théorème 4). Or pour les systèmes hamiltoniens la stabilité asymptotique est impossible, donc la stabilité dans l'approximation linéaire est toujours neutre et insuffisante pour conclure sur la stabilité de la position d'équilibre du problème non linéaire.

Exemple 2. Considérons un écoulement parallèle plan sur le tore

$$\{(x, y), x \bmod X, y \bmod 2\pi\}$$

de champ de vitesses $v = (\sin y, 0)$ parallèle à l'axe x. Ce champ est défini par la fonction de courant $\psi = -\cos y$ et admet pour rotationnel $r = -\cos y$. Le profil des vitesses présente deux points d'inflexion, cependant la fonction de courant s'exprime en fonction de la fonction du rotationnel. Le quotient $\nabla\psi/\nabla\Delta\psi$ est égal à -1. En appliquant le théorème 13 on s'assure que notre écoulement stationnaire est stable si

$$\int_0^{2\pi}\int_0^X (\Delta\varphi)^2\,dx\,dy \geqslant \int_0^{2\pi}\int_0^X (\nabla\varphi)^2\,dx\,dy$$

pour toutes les fonctions φ de période X en x et 2π en y. On établit immédiatement que la dernière inégalité n'a lieu que pour $X \leqslant 2\pi$.

Donc le théorème 13 entraîne la stabilité de l'écoulement stationnaire sinusoïdal dans le cas d'un tore court lorsque la période dans le sens de l'écoulement principal (X) est inférieure à la largeur (2π) de cet écoulement. Par ailleurs on peut immédiatement vérifier que sur un tore long (pour $X > 2\pi$) l'écoulement sinusoïdal est instable *).

Donc dans l'exemple considéré la condition suffisante de stabilité du théorème 13 est aussi une condition nécessaire.

A remarquer qu'en général si la forme quadratique d^2H n'est pas définie positive ou négative, cela ne signifie nullement que l'écoulement correspondant est instable. D'une façon générale la position d'équilibre d'un système hamiltonien peut être stable même si la fonction de Hamilton n'est ni un maximum ni un minimum dans cette position d'équilibre. Le hamiltonien quadratique $H = p_1^2 + q_1^2 - p_2^2 - q_2^2$ en est un exemple simple.

*) Voir par exemple l'article: L. Mechalkine et Y. Sinaï, *Etude de la stabilité de l'écoulement stationnaire d'un système d'équations du mouvement plan d'un fluide visqueux incompressible* (en russe). « Prikladnaïa matematika i mekhanika », n° 6 (1961).

K. Courbure riemannienne d'un groupe de difféomorphismes. L'expression de la courbure d'un groupe de Lie muni d'une métrique invariante unilatérale que nous avons établie dans le point F est également valable pour le groupe SDiff D de difféomorphismes d'un domaine riemannien D. Ce groupe est l'espace de configuration d'un fluide parfait remplissant le domaine D. L'énergie cinétique du fluide définit sur le groupe SDiff D une métrique invariante à droite. Le nombre que l'on obtient en appliquant formellement à ce groupe infini la formule de la courbure d'un groupe de Lie s'appelle naturellement courbure du groupe SDiff D.

Le calcul de la courbure d'un groupe de difféomorphismes n'a été mené jusqu'au bout que dans le cas d'écoulements sur un tore bidimensionnel muni d'une métrique euclidienne. Ce tore s'obtient à partir du plan euclidien $\mathbf{R}^2$ par identification des points dont la différence appartient à un treillis Γ (un sous-groupe discret du plan $\mathbf{R}^2$). Un exemple d'un tel treillis est l'ensemble des points de coordonnées entières. Dans le cas général d'un treillis quelconque Γ le carré qui repose à la base de ce treillis spécial peut être remplacé par un parallélogramme quelconque.

Considérons maintenant l'algèbre de Lie formée par les champs de vecteurs de divergence nulle sur un tore de fonction de courant univoque. Le groupe S_0Diff T^2 correspondant est composé de difféomorphismes laissant invariant le centre de gravité du tore et respectant l'élément d'aire. Il est plongé dans le groupe SDiff T^2 de tous les difféomorphismes respectant l'élément d'aire comme une sous-variété totalement géodésique (i.e. une sous-variété telle que chacune de ses géodésiques est géodésique de la variété ambiante).

L'idée générale de la démonstration est la suivante : si à la date initiale le champ de vitesses du fluide parfait possède une fonction de courant univoque, alors cette fonction le restera tout le temps ; ceci découle de la loi de conservation de l'impulsion.

L'on se propose maintenant d'étudier la courbure du groupe S_0Diff T^2 dans toutes les sections passant par l'unité du groupe (la courbure du groupe SDiff T^2 dans chacune de ces sections est la même, puisque la sous-variété S_0Diff T^2 est totalement géodésique).

Faisons choix d'une orientation sur le plan $\mathbf{R}^2$. Alors les éléments de l'algèbre de Lie du groupe S_0Diff T^2 peuvent être considérés comme des fonctions réelles de valeur moyenne nulle sur le tore (un champ de divergence nulle se déduit d'une telle fonction si elle est regardée comme une fonction de courant). Donc une section dans un plan tangent au groupe S_0Diff T^2 est définie par un couple de fonctions de valeur moyenne nulle sur le tore.

Une telle fonction sera définie par la donnée de ses coefficients de Fourier. Les calculs sur les séries sont commodes à effectuer dans le domaine complexe. Désignons par e_k (k est un point de notre plan euclidien appelé vecteur d'onde) une fonction dont la valeur en un

point x du plan est $e^{i(k,x)}$. Cette fonction définit une fonction sur le tore si elle est Γ-périodique, i.e. si l'addition d'un vecteur du treillis Γ à l'argument x ne modifie pas les valeurs de cette fonction.

En d'autres termes, le produit scalaire (k, x) doit être un multiple de 2π pour tous les $x \in \Gamma$. Tous les vecteurs k appartiennent à un treillis Γ^* sur le plan $\mathbf{R}^2$. Les fonctions e_k, où $k \in \Gamma^*$, forment un système complet dans l'espace des fonctions complexes sur le tore.

On se propose maintenant de complexifier notre algèbre de Lie, le produit scalaire $\langle , \rangle$, le commutateur $[,]$ et l'opération B ainsi que la connexité riemannienne et le tenseur de courbure Ω : toutes ces fonctions seront donc (multi) linéaires dans l'espace vectoriel de l'algèbre complexifiée de Lie. Les fonctions e_k (où $k \in \Gamma^*$, $k \neq 0$) forment une base dans cet espace vectoriel.

Théorème 14. *Les formules du produit scalaire, du commutateur, de l'opération B, de la connexité et la courbure d'une métrique invariante à droite sur le groupe* S_0Diff T^2 *s'écrivent*

$$\langle e_k, e_l\rangle = 0 \text{ pour } k+l \neq 0, \quad \langle e_k, e_{-k}\rangle = k^2 S;$$

$$[e_k, e_l] = (k \wedge l)\, e_{k+l};$$

$$B(e_k, e_l) = b_{k,l} e_{k+l}, \quad \text{où} \quad b_{k,l} = (k \wedge l)\frac{k^2}{(k+l)^2};$$

$$\nabla_{e_k} e_l = d_{l,k+l} e_{k+l}, \quad \text{où} \quad d_{u,v} = \frac{(v \wedge u)(u \cdot v)}{v^2};$$

$R_{k,l,m,n} = 0$, *si* $k+l+m+n \neq 0$; *si par contre* $k+l+m+n = 0$, *alors*

$$R_{k,l,m,n} = (a_{ln}a_{km} - a_{lm}a_{kn})\, S,$$

où $a_{uv} = \dfrac{(u \wedge v)^2}{|u+v|}$.

Dans ces formules S est l'aire du tore, $u \wedge v$ l'aire du parallélogramme construit sur u et v (dans l'orientation choisie de $\mathbf{R}^2$). Les parenthèses représentent le produit scalaire sur le plan, les chevrons— sur l'algèbre de Lie.

La démonstration de ce théorème figure dans l'article cité sous la référence 1 à la page 319.

Les formules mentionnées permettent de calculer une courbure sectionnelle quelconque. Les calculs montrent que cette courbure est négative dans la plupart des sections et positive dans quelques-unes seulement. Considérons un écoulement quelconque du fluide, i.e. une géodésique de notre groupe. En vertu de l'équation de Jacobi la stabilité de cette géodésique est déterminée par les courbures sectionnelles dans tous les 2-plans passant par le vecteur vitesse de la géodésique en chacun de ses points.

Supposons maintenant que l'écoulement considéré est stationnaire. Alors la géodésique est un sous-groupe à un paramètre de notre

groupe. Il résulte qu'en tous les plans passant par le vecteur vitesse de la géodésique en chacun de ses points, les courbures sont égales aux courbures dans les plans correspondants passant par le vecteur vitesse de la géodésique indiquée à la date initiale. (Démonstration: translation à droite dans l'unité du groupe.) Donc la stabilité de l'écoulement stationnaire n'est influencée que par les courbures sectionnelles dans les 2-plans de l'algèbre de Lie qui contiennent le vecteur du champ de vitesses de l'écoulement stationnaire.

Considérons l'exemple simple d'un écoulement stationnaire sinusoïdal parallèle. Il est défini par la fonction de courant

$$\xi = \frac{e_k + e_{-k}}{2}.$$

Soit un vecteur quelconque réel de l'algèbre, $\eta = \sum x_l e_l$ (de sorte que $x_{-l} = \bar{x}_l$). Du théorème 14 on déduit sans peine le

Théorème 15. *La courbure du groupe* $S_0\mathrm{Diff}\, T^2$ *n'est pas positive dans tous les 2-plans contenant la direction de* ξ. *Plus exactement*

$$\langle \Omega(\xi, \eta)\, \xi, \eta \rangle = -\frac{S}{4} \sum_l a_{k,l}^2 \,|\, x_l + x_{l+2k} \,|^2.$$

De cette formule on déduit en particulier que:

1. *la courbure est nulle seulement dans les 2-plans qui sont constitués d'écoulements parallèles de même direction que* ξ, *de sorte que* $[\xi, \eta] = 0$;

2. *dans la section définie par les fonctions de courant* $\xi = \cos kx$ et $\eta = \cos lx$ *la courbure est égale à*

$$K = -\frac{k^2 + l^2}{4S} \sin^2 \alpha \sin^2 \beta,$$

où S *est l'aire du tore,* α *l'angle que fait* k *avec* l, β *l'angle que fait* $k + l$ *avec* $k - l$;

3. *en particulier, la courbure du groupe de difféomorphismes du tore* $\{(x, y) \bmod 2\pi\}$ *dans la section définie par les champs de vitesses* $(\sin y, 0)$ *et* $(0, \sin x)$ *est égale à*

$$K = -\frac{1}{8\pi^2}.$$

L. Discussion. Il est naturel d'attendre de la courbure du groupe de difféomorphismes qu'elle soit liée à la stabilité des géodésiques sur ce groupe (i.e. à la stabilité des écoulements du fluide parfait) comme la courbure d'un groupe fini de Lie l'est avec la stabilité des géodésiques sur ce groupe. Plus exactement, la négativité de la courbure entraîne l'instabilité exponentielle des géodésiques. Ceci étant, le chemin caractéristique (le chemin moyen sur lequel les erreurs sur les conditions initiales grandissent de e fois) est du même ordre que

la grandeur $1/\sqrt{-K}$. Donc la connaissance de la courbure du groupe de difféomorphismes permet de déterminer le temps sur lequel on peut prédire le futur de l'écoulement du fluide parfait d'après le champ de vitesses initial approché sans que l'erreur commise croisse de plusieurs ordres de grandeur.

A souligner que l'instabilité des écoulements d'un fluide parfait n'a pas le même sens que dans le point J : il s'agit de l'instabilité exponentielle du *mouvement du fluide* et non pas de son champ de vitesses. Il peut arriver que l'écoulement stationnaire soit solution stable au sens de Liapounov de l'équation d'Euler, tandis que le mouvement correspondant du fluide est exponentiellement instable. En effet, toute petite variation du champ de vitesses du fluide est susceptible de provoquer une variation du mouvement du fluide qui s'amplifiera exponentiellement. Dans un tel cas (qui correspond à la stabilité de la solution de l'équation d'Euler et à la négativité de la courbure du groupe) on peut se livrer à des prévisions sur le champ de vitesses, mais il est impossible de dire quoi que ce soit sur le mouvement des masses du fluide sans risque de commettre une lourde erreur.

Les formules précédentes de la courbure peuvent être utilisées pour une évaluation grossière de l'intervalle de temps sur lequel il est impossible de se livrer à une prévision météorologique dynamique durable, sous réserve de quelques conventions simplificatrices. On suppose notamment que

1. la terre a la forme d'un tore obtenu par factorisation du plan sur un treillis carré;
2. l'atmosphère est un fluide homogène parfait incompressible bidimensionnel;
3. le mouvement de l'atmosphère est proche d'un « flot alizé » parallèle à l'équateur du tore et de profil de vitesses sinusoïdal.

Pour calculer le chemin caractéristique nous devons évaluer la courbure du groupe $S_0\mathrm{Diff}\ T^2$ dans les sections contenant le « flot alizé » ξ du théorème 15. Ceci étant, nous poserons $T^2 = \{(x, y) \bmod 2\pi\}$, $k = (0, 1)$. En d'autres termes nous considérons des écoulements 2π-périodiques sur le plan (x, y) voisins d'un écoulement stationnaire parallèle à l'axe x et de profil de vitesses sinusoïdal

$$\mathbf{v} = (\sin y,\ 0).$$

Des formules du théorème 15 il découle visiblement que la courbure du groupe $S_0\mathrm{Diff}\ T^2$ dans les plans contenant notre flot alizé $\mathbf{v}$ varie dans l'intervalle

$$-\frac{2}{S} < K < 0, \text{ où } S = 4\pi^2 \text{ est l'aire du tore.}$$

La borne inférieure est une grossière majoration. Cependant il existe de toute évidence une section dans laquelle la courbure est $K = -1/(2S)$ et de nombreuses autres dans lesquelles elle est proche de cette valeur. Pour évaluer le chemin caractéristique nous allons prendre la valeur $K_0 = -1/(2S)$ pour « courbure moyenne ».

Si l'on convient d'effectuer les calculs à partir de cette valeur K_0 de la courbure, on obtient pour le chemin caractéristique

$$s = (\sqrt{-K_0})^{-1} = \sqrt{2S}.$$

La vitesse du mouvement qui correspond sur le groupe à notre flot alizé est égale à $\sqrt{S/2}$ (puisque la moyenne du carré du sinus est égale à $^1/_2$). Donc l'écoulement parcourra le chemin caractéristique en un temps égal à 2. Les particules de courant les plus rapides parcourront pendant ce temps une distance égale à 2, i.e. $1/\pi$ d'un tour complet de tore.

Si donc l'on adopte la valeur approchée proposée plus haut de la courbure moyenne, les erreurs augmenteront de $e^{\pi} \approx 20$ fois le temps que la plus rapide particule effectuera un tour complet. Si l'on suppose que la vitesse maximale du flot alizé est de 100 km/h, il faudra alors 400 h pour effectuer un tour complet, soit moins de trois semaines.

Donc si à la date initiale on connaît le temps avec une petite erreur ε, au bout de n mois l'erreur commise sur les prévisions sera de

$$10^{kn}\varepsilon, \quad \text{où} \quad k \approx \frac{30 \cdot 24}{400} \pi \log e \approx 2{,}5.$$

Pour fixer les idées, si l'on prédit la météo deux mois à l'avance, il faut avoir en réserve 5 décimales. Pratiquement cela veut dire qu'il est impossible de prévoir la météo pour une telle date à l'avance.

Certes la majoration adoptée est grossière et le modèle considéré est très simplifié. Quant au choix de la « courbure moyenne » il a besoin d'être justifié.

STRUCTURE SYMPLECTIQUE SUR DES VARIÉTÉS ALGÉBRIQUES

Les variétés symplectiques de la mécanique classique sont souvent les espaces de phases de systèmes mécaniques lagrangiens, i.e. les fibrés cotangents aux espaces de configuration.

La géométrie algébrique nous offre un tout autre exemple de variétés symplectiques.

Par exemple, toute variété algébrique complexe différentiable (définie par un système d'équations polynomiales sur l'espace projectif complexe) possède une structure symplectique naturelle.

La construction d'une structure symplectique sur une variété algébrique est fondée sur le fait que l'espace projectif complexe possède lui-même une remarquable structure symplectique, plus exactement la partie imaginaire de sa structure hermitienne.

A. Structure hermitienne de l'espace projectif complexe. On rappelle que l'espace projectif complexe CP^n est la variété formée par toutes les droites complexes passant par le point O de l'espace vectoriel complexe C^{n+1}. Pour construire une structure symplectique sur l'espace projectif complexe CP^n nous allons nous servir de la structure hermitienne de l'espace vectoriel correspondant C^{n+1}.

On rappelle que le *produit scalaire hermitien* (ou structure hermitienne) dans l'espace vectoriel complexe est par définition une fonction complexe d'un couple de vecteurs qui : 1) est linéaire en le premier et antilinéaire en le second argument, 2) prend la valeur conjuguée complexe si l'on permute ses arguments et 3) se transforme en une forme quadratique réelle définie positive si les arguments sont égaux entre eux, i.e. :

$$\langle \lambda\xi, \eta\rangle = \lambda\langle \xi, \eta\rangle, \quad \langle \eta, \xi\rangle = \overline{\langle \xi, \eta\rangle}, \quad \langle \xi, \xi\rangle > 0$$

pour $\xi \neq 0$.

Un exemple de produit scalaire hermitien est

$$\langle \xi, \eta\rangle = \sum \xi_k \bar{\eta}_k, \qquad (1)$$

où ξ_k et η_k sont les composantes des vecteurs ξ et η dans une certaine base.

La base par rapport à laquelle le produit scalaire hermitien prend la forme (1) existe toujours et s'appelle base orthonormée hermitienne.

Les parties réelle et imaginaire du produit scalaire hermitien sont des formes bilinéaires réelles. La première est symétrique, la seconde est antisymétri-

que et toutes deux sont non dégénérées :

$$\langle \xi, \eta \rangle = (\xi, \eta) + i[\xi, \eta], \quad (\xi, \eta) = (\eta, \xi), \quad [\xi, \eta] = -[\eta, \xi].$$

La forme quadratique (ξ, ξ) est définie positive.

Donc la structure hermitienne $\langle , \rangle$ définit une structure euclidienne $(,)$ et une structure symplectique $[,]$ sur l'espace vectoriel complexe. Ces deux structures sont liées à la structure complexe par la relation

$$[\xi, \eta] = (\xi, i\eta).$$

Définissons maintenant une métrique riemannienne sur l'espace projectif complexe. Considérons à cet effet sur l'espace vectoriel correspondant $\mathbf{C}^{n+1}$ la sphère unité

$$S^{2n-1} = \{z \in \mathbf{C}^{n+1} : \langle z, z \rangle = 1\}.$$

Cette sphère hérite sa structure riemannienne de l'espace $\mathbf{C}^{n+1}$. Toute droite complexe coupe cette sphère suivant un grand cercle.

Définition. On appelle *distance* de deux points de l'espace projectif complexe la distance de deux cercles correspondants sur la sphère unité.

Remarquons que ces deux cercles sont parallèles en ce sens que la distance d'un point quelconque de l'un des cercles à l'autre est la même (démonstration : le produit de z par $e^{i\varphi}$ conserve la métrique sur la sphère). Ceci permet aussitôt d'écrire la formule explicite (2) pour la métrique riemannienne sur l'espace projectif complexe donnant la distance définie plus haut.

En effet, soit p l'application

$$p : \mathbf{C}^{n+1} \setminus 0 \to \mathbf{C}P^n,$$

qui à tout point $z \neq 0$ de $\mathbf{C}^{n+1}$ associe la droite complexe qui passe par O et z.

Tout vecteur ζ tangent à $\mathbf{C}P^n$ au point pz peut être (de plusieurs manières) représenté par l'image du vecteur appliqué en z ; ceci étant,

$$\zeta = p_* \xi, \quad \xi \in T\mathbf{C}_z^{n+1}.$$

Théorème. *Le carré du module du vecteur ζ dans la métrique riemannienne définie plus haut est donné par la formule*

$$ds^2(\zeta) = \frac{\langle \xi, \xi \rangle \langle z, z \rangle - \langle \xi, z \rangle \langle z, \xi \rangle}{\langle z, z \rangle^2}. \tag{2}$$

Démonstration. Supposons tout d'abord que le point z est situé sur la sphère unité S^{2n-1}.

Développons le vecteur ξ en deux composantes : l'une sur la droite complexe de vecteur directeur z, l'autre dans la direction H-orthogonale (au sens du produit scalaire hermitien). Remarquons que la H-orthogonalité au vecteur z équivaut à la E-orthogonalité aux vecteurs z et iz. Le vecteur z est le vecteur directeur de la normale euclidienne à la sphère S^{2n-1} au point z. Le vecteur iz est le vecteur directeur de la tangente au cercle suivant lequel la sphère coupe la droite complexe passant par le point z. Donc la composante η du vecteur ξ qui est H-orthogonale au vecteur z est tangente à la sphère S^{2n-1} et est E-orthogonale au cercle d'intersection de la sphère et de la droite pz.

En vertu de la définition de la métrique riemannienne sur CP^n, le carré riemannien du module du vecteur ζ est égal au carré euclidien du module de la composante η du vecteur ξ qui est H-orthogonale au vecteur z.

Calculons la composante η du vecteur ξ. Ecrivons notre développement sous la forme

$$\xi = cz + \eta, \quad \text{où} \quad \langle \eta, z \rangle = 0.$$

En multipliant par z au sens de la métrique hermitienne on trouve

$$\langle \xi, z \rangle = c \langle z, z \rangle,$$

donc

$$\eta = \frac{\langle z, z \rangle \xi - \langle \xi, z \rangle z}{\langle z, z \rangle}$$

et

$$\langle \eta, \eta \rangle = \frac{\langle z, z \rangle \langle \xi, \xi \rangle - \langle \xi, z \rangle \langle z, \xi \rangle}{\langle z, z \rangle}.$$

Ce qui démontre la formule (2) pour les points z de la sphère unité. Le cas général se ramène au cas étudié par l'homothétie $z \mapsto z/|z|$. Le théorème est démontré.

On remarquera que notre construction permet de définir non seulement la structure euclidienne (2) mais aussi une structure hermitienne sur tout espace tangent à CP^n.

Considérons en effet le H-orthocomplément H de la direction du vecteur z par rapport à l'espace TC_z^{n+1}, où $z \in S^{2n-1}$. L'application $p_*: H \to T(CP^n)_{pz}$ est un isomorphisme (nous l'avons montré plus haut) de H sur l'espace tangent à CP^n en pz auquel il confère la structure hermitienne de H.

Il est clair que le carré scalaire défini par cette structure hermitienne est donné par la formule (2). Donc la formule du produit scalaire hermitien dans l'espace tangent à CP^n peut s'écrire sans calculs supplémentaires:

$$\langle \zeta_1, \zeta_2 \rangle = \frac{\langle \xi_1, \xi_2 \rangle \langle z, z \rangle - \langle \xi_1, z \rangle \langle z, \xi_2 \rangle}{\langle z, z \rangle^2} \qquad (3)$$

quels que soient ξ_1 et ξ_2 de TC_z^{n+1} tels que $p_* \xi_k = \zeta_k \in T(CP^n)_{pz}$.

A remarquer que dans la formule (3) le point z n'est pas nécessairement situé sur la sphère unité.

Les structures euclidienne (2) et hermitienne (3) construites sur les espaces tangents à CP^n ne sont pas invariantes par toutes les transformations projectives de la variété CP^n mais seulement par celles qui sont induites par des transformations linéaires unitaires (respectant la structure hermitienne) de l'espace vectoriel C^{n+1}.

B. Structure symplectique de l'espace projectif complexe. Considérons la partie imaginaire de la forme hermitienne (3) prise avec le coefficient $-1/\pi$ (la raison de ce choix est expliquée dans l'exercice 1 à la fin de cet Appendice)

$$\Omega(\zeta_1, \zeta_2) = -\frac{1}{\pi} \operatorname{Im} \langle \zeta_1, \zeta_2 \rangle. \qquad (4)$$

Comme la partie imaginaire de toute forme hermitienne, la forme bilinéaire réelle Ω est antisymétrique et non dégénérée sur tout espace tangent à l'espace projectif complexe.

Théorème. *La 2-forme différentielle Ω définit une structure symplectique sur l'espace projectif complexe.*

Démonstration. Il nous suffit de démontrer que la forme Ω est fermée.

Considérons la dérivée extérieure $d\Omega$ de la forme Ω. Cette 3-forme différentielle sur la variété CP^n est invariante par les applications induites par les transformations unitaires de l'espace C^{n+1}. Il s'ensuit qu'elle est nulle.

En effet, considérons dans l'espace tangent à CP^n en un point quelconque z une C-base H-orthonormée $e_1, \ldots, e_n$. Les vecteurs $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ forment une R-base E-orthonormée. Montrons que la forme $d\Omega$ est nulle sur un triplet de vecteurs quelconques de cette R-base. (Nous supposons que $n > 1$, car pour $n = 1$ il n'y a rien à démontrer.)

Remarquons que dans tout triplet de vecteurs de la R-base il en existe au moins un qui est hermito-orthogonal aux deux autres. Désignons ce vecteur par e. Il est aisé de construire une transformation unitaire de l'espace C^{n+1} induisant dans CP^n un mouvement laissant fixe le point considéré z et le H-orthocomplément de e et modifiant le sens du vecteur e.

La forme $d\Omega$ prend sur le triplet de vecteurs e, f, g la même valeur que sur le triplet $-e, f, g$ par suite de son invariance et par conséquent est nulle. Le théorème est démontré.

Remarque. Voyons un autre procédé de construction de la même structure symplectique sur l'espace projectif complexe. Considérons pour cela les petites oscillations d'un pendule mathématique dont l'espace de configuration est de dimension $(n + 1)$. Utilisons maintenant l'intégrale d'énergie pour abaisser d'une unité le nombre de degrés de liberté du système. L'espace de phases obtenu par cette opération est la variété CP^n, et la structure symplectique définie sur elle coïncide à un facteur près avec la forme Ω.

Un troisième procédé de construction d'une structure symplectique sur CP^n consiste à représenter cet espace comme une orbite de la représentation coadjointe d'un groupe de Lie ; sur chaque orbite il existe toujours une structure symplectique standard (voir Appendice 2, point A). Comme groupe de Lie on peut prendre le groupe des opérateurs unitaires (respectant la métrique hermitienne) dans l'espace complexe C^{n+1}. Les orbites de la représentation coadjointe sont dans ce cas les mêmes que celles de la représentation adjointe. Dans cette dernière, l'opérateur de réflexion dans un hyperplan (qui change le signe de la première coordonnée et pas celui des autres) admet CP^n pour orbite, car il est univoquement défini par la droite complexe orthogonale en 0 à ce plan.

C. Structures symplectiques sur des variétés algébriques projectives. Nous obtenons maintenant une structure symplectique sur toute sous-variété complexe M de l'espace projectif complexe. Plus exactement, soit $J: M \to CP^n$ le plongement de la variété complexe M dans l'espace projectif complexe. Les structures riemannienne, hermitienne et symplectique définies sur l'espace projectif induisent des structures correspondantes sur M. Par exemple, une structure

symplectique sur M est définie par la formule

$$\Omega_M = j^*\Omega.$$

Théorème. *La forme différentielle Ω_M définit une structure symplectique sur la variété M.*

D é m o n s t r a t i o n. La non-dégénérescence de la 2-forme Ω_M découle du fait que M est une sous-variété complexe. En effet, la forme quadratique

$$(\xi, \xi) = \Omega_M (\xi, i\xi)$$

est définie positive sur TM_x (elle est induite par une métrique riemannienne sur CP^n). Par suite la forme bilinéaire $(\xi, \eta) = \Omega_M (\xi, i\eta)$ est non dégénérée. Donc est non dégénérée et la forme Ω_M. La fermeture de la forme Ω_M découle de celle de Ω. Le théorème est démontré.

R e m a r q u e. Sur les espaces tangents à l'espace projectif complexe et donc à ses sous-variétés complexes nous avons défini une structure hermitienne dont la partie imaginaire est une structure symplectique.

Une variété complexe munie d'une métrique hermitienne dont la partie imaginaire est une forme fermée (i.e. une structure symplectique) s'appelle *variété kählérienne* et sa métrique hermitienne *métrique kählérienne*. De nombreux importants résultats ont été obtenus en géométrie des variétés kählériennes: ces dernières possèdent de remarquables propriétés topologiques (voir par exemple André W e i l, *Introduction à l'étude des variétés kählériennes*, Hermann, Paris, 1958).

Les variétés symplectiques connues n'admettent pas toutes une structure kählérienne.

E x e r c i c e 1. Calculer la structure symplectique Ω dans la carte affine $w_k = z_1 : z_0$ de la droite projective CP^1.

Réponse. $\Omega = \frac{1}{\pi} \frac{dx \wedge dy}{(1+x^2+y^2)^2}$, où $w = x + iy$.

Dans la définition de la forme Ω, le coefficient est choisi de manière que la droite complexe ait l'orientation usuelle $(dx \wedge dy)$ et que l'intégrale de la forme Ω sur toute la droite projective soit égale à l'unité.

E x e r c i c e 2. Montrer que dans la carte affine $w_k = z_k z_0^{-1}$ $(k = 1, \ldots, n)$ de l'espace projectif $CP^n = \{(z_0, z_1, \ldots, z_n)\}$ la structure symplectique Ω est définie par la formule

$$\Omega = \frac{i}{2\pi} \frac{\sum\limits_{0 \leq k < l \leq n} (w_k\, dw_l - w_l\, dw_k) \wedge (\overline{w}_k\, d\overline{w}_l - \overline{w}_l\, d\overline{w}_k)}{\sum\limits_{k=0}^{n} (w_k \overline{w}_k)^2}.$$

R e m a r q u e. Dans l'espace complexe les formes différentielles à valeurs complexes (par exemple dw_k et $d\overline{w}_k$) se définissent comme des fonctions linéaires

complexes de vecteurs tangents; si $w_k = x_k + iy_k$, alors

$$dw_k = dx_k + i\,dy_k, \quad d\bar{w}_k = dx_k - i\,dy_k.$$

L'espace de telles formes sur $\mathbf{C}^n$ possède une dimension complexe $2n$; une $\mathbf{C}$-base peut par exemple être constituée par les $2n$ formes dw_k, $d\bar{w}_k$ ($k = 1, \ldots, n$) ou par les $2n$ formes dx_k, dy_k.

Le produit extérieur se définit comme d'habitude et obéit aux règles ordinaires. Par exemple

$$dw \wedge d\bar{w} = (dx + i\,dy) \wedge (dx - i\,dy) = -2i\,dx \wedge dy.$$

Soit f une fonction réelle différentiable sur $\mathbf{C}^n$ (à valeurs en général complexes). Exemple: $|w|^2 = \sum w_k \bar{w}_k$. La différentielle de la fonction f est une 1-forme complexe. Donc elle peut être décomposée suivant la base dw_k, $d\bar{w}_k$. Les coefficients de ce développement sont appelés dérivées partielles « par rapport à w_k » et « par rapport à $\bar{w}_k$ »:

$$df = \frac{\partial f}{\partial w}\,dw + \frac{\partial f}{\partial \bar{w}}\,d\bar{w}.$$

Dans le calcul des dérivées extérieures, il est commode de distinguer la dérivation d' par rapport à w et la dérivation d'' par rapport à $\bar{w}$, de sorte que $d = d' + d''$.

Pour la fonction f par exemple

$$d'f = \frac{\partial f}{\partial w}\,dw, \quad d''f = \frac{\partial f}{\partial \bar{w}}\,d\bar{w}.$$

Pour la 1-forme différentielle

$$\omega = \sum a_k\,dw_k + b_k\,d\bar{w}_k$$

les opérateurs d' et d'' se définissent de façon analogue:

$$d'\omega = \sum d'a_k \wedge dw_k + d'b_k \wedge d\bar{w}_k,$$

$$d''\omega = \sum d''a_k \wedge dw_k + d''b_k \wedge d\bar{w}_k.$$

Exercice 3. Montrer que dans la carte affine $w_k = z_k z_0^{-1}$ de l'espace projectif $\mathbf{CP}^n$ la structure symplectique Ω est définie par la formule

$$\Omega = \frac{i}{2\pi}\,d'd''\ln \sum_{k=0}^{n} |w_k|^2.$$

APPENDICE 4

STRUCTURES DE CONTACT

Il n'existe pas de structure symplectique sur une variété de dimension impaire. Son analogue pour les variétés de dimension impaire est une structure un peu moins symétrique, mais jouissant également de propriétés remarquables : la structure de contact.

En mécanique les structures symplectiques ont pour origine les espaces de phases (i.e. les fibrés cotangents aux variétés de configuration) sur lesquels existe toujours une structure symplectique canonique. Les structures de contact ont pour source les variétés des éléments de contact des espaces de configuration.

On appelle élément de contact à une variété différentiable de dimension n en un point quelconque un $(n - 1)$-plan tangent à cette variété en ce point (i.e. un sous-espace vectoriel de dimension $(n - 1)$ de l'espace de dimension n tangent en ce point).

L'ensemble de tous les éléments de contact d'une variété de dimension n possède une structure naturelle de variété différentiable de dimension $(2n - 1)$. Il s'avère que sur cette variété de dimension impaire il existe encore une remarquable structure de contact (que nous décrirons plus bas).

La variété des éléments de contact d'une variété riemannienne de dimension n est étroitement liée avec la variété de dimension $(2n - 1)$ des vecteurs unitaires tangents de cette variété riemannienne de dimension n ou avec la variété de dimension $(2n - 1)$ de niveau d'énergie d'un point matériel se déplaçant par inertie sur une variété riemannienne. Dans ces variétés de dimension $(2n - 1)$ les structures de contact sont étroitement liées avec la structure symplectique de l'espace de phases de dimension $2n$ du point (i.e. avec la structure du fibré cotangent à la variété riemannienne de départ).

A. Définition de la structure de contact.

D é f i n i t i o n. Une *structure de contact* sur une variété est un champ différentiable d'hyperplans tangents *) vérifiant une

*) On appelle hyperplan d'un espace vectoriel de dimension n un sous-espace de dimension $(n - 1)$ (i. e. l'ensemble de niveau nul d'une fonction linéaire non identiquement nulle).

Un hyperplan tangent est un hyperplan d'un espace tangent.

condition de non-dégénérescence qui sera formulée plus bas. Pour formuler cette condition voyons quelle est la forme du champ d'hyperplans au voisinage d'un point d'une variété de dimension N.

Exemple. Supposons $N = 2$. La variété est alors une surface et le champ d'hyperplans un champ de droites. Au voisinage d'un point l'allure d'un tel champ est simple et toujours la même, à savoir celle d'un champ de tangentes à une famille de droites parallèles sur le plan. Plus exactement, l'un des principaux résultats de la théorie locale des équations différentielles ordinaires consiste en la possibilité de transformer tout champ différentiable de droites tangentes sur une variété en un champ de tangentes à une famille de droites parallèles de l'espace euclidien par un difféomorphisme dans un voisinage suffisamment petit de tout point de cette variété.

Si $N > 2$, l'hyperplan n'est plus une droite et la situation se complique. Par exemple, un champ de 2-plans tangents de l'espace euclidien ordinaire ne peut pas toujours être appliqué par un difféomorphisme sur un champ de plans parallèles. En effet, il existe des champs de plans tangents pour lesquels il est impossible de tracer une « surface intégrale », i.e. une surface qui admette le plan tangent requis en chacun de ses points.

La condition de non-dégénérescence du champ d'hyperplans qui participe à la définition de la structure de contact se traduit par un éloignement maximum du champ d'hyperplans par rapport au champ de tangentes à la famille d'hypersurfaces. Quelques constructions et des calculs nous seront nécessaires pour mesurer cet éloignement et en général pour nous assurer de l'existence de champs ne possédant pas d'hypersurfaces intégrales *).

B. Condition d'intégrabilité de Frobenius. Considérons un point quelconque d'une variété de dimension N et essayons de construire une surface passant par ce point et tangente en chacun de ses points à un champ donné de $(N - 1)$-plans (surface intégrale).

A cet effet considérons au voisinage du point envisagé un système de coordonnées tel qu'une surface arithmétique soit tangente en ce point à un plan du champ. Nous dirons que ce plan est horizontal et que l'axe de coordonnées qui n'y est pas contenu est vertical.

Construction de la surface intégrale. La surface intégrale, si elle existe, est le graphe d'une fonction de $N - 1$ variables au voisinage de l'origine des coordonnées. Pour la construire nous pouvons considérer un chemin différentiable quelconque sur le plan horizontal. Alors les verticales issues de ce chemin engendrent une 2-surface (un cylindre) sur laquelle notre champ de plans

*) A partir d'ici nous omettrons le préfixe « hyper ». Si l'on veut, l'on pourra considérer que l'on se trouve dans un espace de dimension trois et les hypersurfaces seront des surfaces ordinaires. Le cas multidimensionnel est analogue au cas tridimensionnel.

découpe un champ de droites tangentes. La surface intégrale (si elle existe) coupe le cylindre suivant une courbe intégrale du champ de droites, issue de l'origine des coordonnées. Une telle courbe intégrale existe toujours que la surface intégrale existe ou non. Donc nous pouvons construire une surface intégrale au-dessus du plan horizontal en nous déplaçant sur les courbes différentiables de ce dernier.

Ensuite, pour que toutes les courbes intégrales engendrent une surface intégrale différentiable, il faut que notre construction ne dépende pas du chemin mais seulement du point final.

En particulier, si l'on fait un tour complet sur un chemin fermé au voisinage de l'origine des coordonnées sur le plan horizontal, la courbe intégrale doit se fermer sur le cylindre.

Il est aisé de construire des champs de plans pour lesquels cette fermeture n'a pas lieu et par conséquent la surface intégrale n'existe pas. De tels champs sont dits *non intégrables.*

Exemple de champ de plans non intégrable. Introduisons les notations suivantes pour définir un champ de plans et mesurer son écart par rapport à la fermeture.

Remarquons tout d'abord qu'un champ d'hyperplans peut être donné localement par une 1-forme différentielle. En effet, un plan tangent définit une 1-forme à la multiplication près par une constante non nulle. Choisissons cette constante de telle sorte que la valeur de la 1-forme soit égale à 1 sur les vecteurs de base tangents verticaux.

Cette condition peut être réalisée dans un voisinage de l'origine des coordonnées, puisque le plan du champ ne contient pas de verticale en 0. Cette condition définit univoquement la forme (d'après le champ des plans).

Dans l'espace ordinaire, un champ de plans ne possédant pas de surfaces intégrales peut être défini, par exemple, par la 1-forme

$$\omega = x\,dy + dz,$$

où x et y sont les coordonnées horizontales et z la coordonnée verticale. On démontre plus bas que ce champ n'est pas intégrable.

Construction d'une 2-forme mesurant la non-intégrabilité. On peut mesurer le degré de non-intégrabilité à l'aide de la forme qui définit le champ moyennant la construction suivante (fig. 236).

Considérons un couple de vecteurs issus de l'origine des coordonnées et situés dans le plan horizontal de notre système de coordonnées. Construisons un parallélogramme sur ces vecteurs. Nous obtenons deux chemins pour aller de l'origine des coordonnées au sommet opposé. Au-dessus de chacun de ces chemins on peut construire une courbe intégrale (à deux segments). Nous obtenons ainsi, d'une façon générale, deux points au-dessus du sommet opposé à l'origine des coordonnées. La différence de cotes de ces deux points est une fonction

du couple de vecteurs envisagés. Cette fonction est antisymétrique et nulle si l'un des vecteurs l'est. Donc la partie linéaire du développement taylorien de cette fonction est nulle en 0, et sa partie quadratique est une forme antisymétrique bilinéaire sur le plan horizontal.

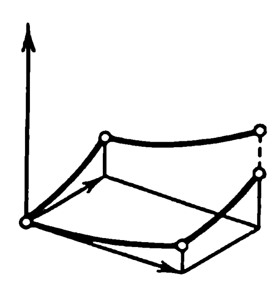


Fig. 236. Courbes intégrales construites d'après un champ de plans non intégrable.

Si le champ est intégrable, la 2-forme obtenue est nulle. On peut donc considérer cette 2-forme comme la mesure de la non-intégrabilité du champ.

En se servant du système de coordonnées considéré on peut identifier le plan arithmétique horizontal au plan du champ qui passe par l'origine des coordonnées. Donc notre construction introduit une 2-forme sur un plan même du champ.

La 2-forme est intrinsèquement définie. Cette 2-forme a été construite à l'aide des coordonnées. Cependant sur un couple de vecteurs tangents elle prend une valeur qui ne dépend pas du système de coordonnées mais uniquement de la 1-forme qui a servi à définir le champ.

Pour s'en assurer il suffit de démontrer le

Théorème. *La 2-forme définie ci-dessus sur l'espace des zéros de la 1-forme ω coïncide avec la dérivée extérieure de cette dernière* : $d\omega\,|_{\omega=0}$.

Démonstration. Montrons que la différence de cotes des deux points de la construction précédente est confondue avec l'intégrale de la 1-forme ω étendue aux quatre côtés du parallélogramme aux infiniment petits près du troisième ordre par rapport aux côtés du parallélogramme.

Remarquons à ces fins que la cote d'une courbe intégrale au-dessus d'un chemin quelconque de longueur ε issu de l'origine des coordonnées est de l'ordre de ε^2, puisque en l'origine des coordonnées le plan du champ est horizontal. Donc les intégrales de la 2-forme $d\omega$ étendues aux quatre aires verticales construites sur les côtés du parallélogramme et interceptées par les courbes intégrales et le plan horizontal sont de l'ordre de ε^2 si les côtés sont de l'ordre de ε.

Les intégrales de la forme ω étendues aux courbes intégrales sont nulles. Donc en vertu de la formule de Stokes l'accroissement de la cote d'une courbe intégrale située au-dessus de l'un quelconque des côtés du parallélogramme est égal à l'intégrale de la 1-forme ω le long de ce côté aux infiniment petits près du troisième ordre.

A présent le théorème découle immédiatement de la définition de la dérivée extérieure.

Il subsiste encore un arbitraire dans le choix de la 1-forme ω qui nous a servi à construire la 2-forme. Plus exactement, la forme ω est définie par le champ de plans à la multiplication près par une

fonction f nulle part nulle. En d'autres termes, nous aurions pu commencer directement par la forme $f\omega$. Nous serions alors arrivés à la 2-forme

$$d f\omega = f\, d\omega + df \wedge \omega,$$

qui diffère de la 2-forme $d\omega$ par la multiplication par le nombre $f(0)$ non nul.

Donc la *2-forme construite sur un plan du champ est intrinsèquement définie à la multiplication près par une constante non nulle.*

Condition d'intégrabilité d'un champ de plans.

Théorème. *Une condition nécessaire et suffisante pour qu'un champ d'hyperplans soit intégrable est que la 2-forme précédente soit nulle.*

La condition est nécessaire: ceci découle immédiatement de la construction de la 2-forme. On démontre qu'elle est suffisante en reprenant exactement les mêmes raisonnements qui nous ont permis de prouver la commutativité des flots dont est nul le crochet de Poisson des champs de vitesses. On peut tout simplement se référer à cette commutativité et l'appliquer aux courbes intégrales engendrées au-dessus des axes de coordonnées dans le plan horizontal.

Théorème. *La condition d'intégrabilité du champ de plans*

$$d\omega = 0 \textit{ pour } \omega = 0$$

équivaut à la condition suivante de Frobenius

$$\omega \wedge d\omega = 0.$$

Démonstration. Considérons la valeur de la 3-forme $\omega \wedge d\omega$ sur trois vecteurs de base distincts quelconques. Un seul d'entre eux est susceptible d'être vertical. Donc de tous les termes qui figurent dans la définition de la valeur du produit extérieur sur trois vecteurs ne peut être non nul que celui qui est égal au produit de la valeur de la forme ω sur le vecteur vertical par la valeur de la forme $d\omega$ sur le couple de vecteurs horizontaux. Si le champ défini par la forme est intégrable, alors le deuxième terme du produit est nul et notre 3-forme est nulle sur tous les triplets de vecteurs.

Inversement, si la 3-forme est nulle sur des vecteurs quelconques, elle est nulle pour tout triplet de vecteurs de base dont l'un est vertical et les deux autres horizontaux. La valeur de la 3-forme sur un tel triplet est égale au produit de la valeur de ω sur le vecteur vertical par la valeur de $d\omega$ sur le couple de vecteurs horizontaux. Le premier terme du produit n'est pas nul, donc le second est nul et la forme $d\omega$ est nulle sur le plan du champ, c.q.f.d.

C. Champs d'hyperplans non dégénérés.

Définition. Un champ d'hyperplans est *non dégénéré en un point* si le rang de la 2-forme $d\omega\ |_{\omega=0}$ définie sur le plan du champ passant par ce point est égal à la dimension du plan.

Cela signifie que pour tout vecteur non nul de notre plan il doit exister un autre vecteur de ce plan tel que la valeur de la 2-forme sur ce couple de vecteurs soit non nulle.

D é f i n i t i o n. Un champ de plans est *non dégénéré sur une variété* s'il est non dégénéré en chaque point de cette variété.

Remarquons que sur une variété de dimension paire, il ne peut exister de champ d'hyperplans non dégénéré. En effet, sur une telle variété tout hyperplan est de dimension impaire, or le rang de toute forme bilinéaire antisymétrique définie sur un espace de dimension impaire est inférieur à la dimension de cet espace (voir § 44).

Il existe par contre des champs de plans non dégénérés sur des variétés de dimension impaire.

E x e m p l e. Considérons un espace euclidien de dimension $(2m + 1)$ muni des coordonnées x, y, z (où x et y sont des vecteurs d'espaces de dimension m et z un nombre). La 1-forme

$$\omega = x\,dy + dz$$

définit un champ d'hyperplans. L'équation du plan passant par l'origine des coordonnées est $dz = 0$. On peut prendre x et y pour coordonnées dans cet hyperplan. Donc notre 2-forme s'écrira

$$d\omega\,|_{\omega=0} = dx \wedge dy = dx_1 \wedge dy_1 + \ldots + dx_m \wedge dy_m.$$

Le rang de cette forme est égal à $2m$, donc notre champ est non dégénéré en l'origine des coordonnées et partant en son voisinage (en réalité ce champ de plans est non dégénéré en tous les points de l'espace).

Maintenant nous sommes enfin en état de définir une structure de contact sur une variété: *une structure de contact sur une variété est un champ non dégénéré d'hyperplans tangents.*

D. Variété des éléments de contact. Le terme « structure de contact » s'explique par le fait que cette structure existe toujours sur la variété des éléments de contact d'une variété différentiable.

Soit une variété différentiable de dimension n.

D é f i n i t i o n. Un hyperplan (de dimension $(n - 1)$) tangent à la variété en un quelconque de ses points est appelé *élément de contact* et ce point, *point de contact.*

L'ensemble de tous les éléments de contact d'une variété de dimension n possède une structure de variété différentiable de dimension $(2n - 1)$.

En effet, l'ensemble de tous les éléments de contact ayant un point de contact fixe est l'ensemble de tous les sous-espaces de dimension $(n - 1)$ d'un espace vectoriel de dimension n, i.e. l'espace projectif de dimension $(n - 1)$. On définit un élément de contact par la donnée de n coordonnées du point de contact et de $n - 1$ coordonnées d'un point de l'espace projectif de dimension $(n - 1)$, au total $2n - 1$ coordonnées.

La variété de tous les éléments de contact d'une variété de dimension n est l'espace ayant pour base notre variété de dimension n et pour fibre l'espace projectif de dimension $(n - 1)$.

Théorème. *Le fibré des éléments de contact est le projectivisé* *) *du fibré cotangent: on peut l'obtenir à partir du fibré cotangent en remplaçant chaque espace vectoriel cotangent de dimension n par l'espace projectif de dimension* (n — 1) (*dont chaque point est une droite passant par l'origine des coordonnées dans l'espace cotangent*).

En effet, un élément de contact est donné par une 1-forme sur l'espace cotangent pour laquelle cet élément est l'ensemble des zéros. Cette forme n'est pas identiquement nulle et est définie à la multiplication près par une constante non nulle.

Or toute forme sur un espace tangent est un vecteur de l'espace cotangent. Donc une forme non nulle définie à la multiplication près par une constante non nulle sur un espace tangent est un vecteur non nul de l'espace cotangent, défini à la multiplication près par une constante non nulle, i.e. est un point du projectivisé de l'espace cotangent.

S t r u c t u r e d e c o n t a c t s u r u n e v a r i é t é d' é l é m e n t s d e c o n t a c t.

Il existe un remarquable hyperplan dans tout espace cotangent à une variété d'éléments de contact. On l'appelle *hyperplan de contact* et on le définit comme suit.

Fixons un point d'une variété d'éléments de contact de dimension ($2n$ — 1) sur une variété de dimension n. Nous pouvons considérer ce point comme un (n — 1)-plan tangent à la variété initiale de dimension n.

D é f i n i t i o n. On dit qu'un vecteur tangent à une variété d'éléments de contact en un point fixé appartient à un *hyperplan de contact* si sa projection sur la variété de dimension n est contenue dans ce (n — 1)-plan qui est précisément le point fixé de la variété des éléments de contact.

En d'autres termes, *le déplacement d'un élément de contact est tangent à l'hyperplan de contact si la vitesse du point de tangence appartient à cet élément de contact*; *cet élément peut tourner à loisir*.

E x e m p l e. Considérons une sous-variété quelconque de notre variété de dimension n et tous les (n — 1)-plans qui lui sont tangents (i.e. les éléments de contact). Tous ces éléments de contact forment une sous-variété différentiable dans la variété de dimension ($2n$ — 1) de tous les éléments de contact. La dimension de cette sous-variété est égale à (n — 1) *quelle que soit la dimension de la sous-variété initiale* (qui peut être de dimension (n — 1) ou inférieure à (n — 1) et être même une courbe ou un point).

Cette sous-variété de la variété de dimension ($2n$ — 1) de tous les éléments de contact est tangente en chacun de ses points aux hyperplans de contact (par définition de l'hyperplan de contact). Donc *le champ des* ($2n$ — 2)-*hyperplans de contact admet des variétés intégrales de dimension* (n — 1).

*) L'ensemble des vecteurs non nuls d'un espace vectoriel, définis à la multiplication près par une constante non nulle, s'appelle espace projectivisé.

Exercice. Est-ce que ce champ d'hyperplans possède des variétés intégrales de dimension supérieure à $(n - 1)$?

Réponse. Non.

Exercice. Peut-on définir un champ d'hyperplans de contact par une 1-forme différentielle sur la variété de tous les éléments de contact?

Réponse. Non, même si la variété initiale de dimension n est un espace euclidien (par exemple, le plan ordinaire).

Plus bas nous démontrerons que *le champ d'hyperplans de contact sur la variété de dimension* $(2n - 1)$ *de tous les éléments de contact de la variété de dimension n est non dégénéré.*

La démonstration est basée sur la structure symplectique du fibré cotangent.

En effet, la variété des éléments de contact est liée par une simple construction avec l'espace du fibré cotangent (dont les projectivisés des fibres forment la variété des éléments de contact). En outre, la non-dégénérescence du champ des plans de contact du projectivisé du fibré est étroitement liée à la non-dégénérescence de la 2-forme qui définit la structure symplectique du fibré cotangent.

La construction dont il est question sera effectuée ultérieurement dans un cas plus général. Plus exactement, étant donné une variété de dimension impaire munie d'une structure de contact, on peut construire sa « symplectisée », i.e. une variété symplectique dont la dimension est d'une unité supérieure. Ces deux variétés — la variété de contact de dimension impaire et la variété symplectique de dimension paire — sont liées entre elles comme la variété des éléments de contact avec sa structure de contact ou encore comme le fibré cotangent avec sa structure symplectique.

E. Symplectisation d'une variété de contact. Considérons une variété de contact arbitraire, i.e. une variété de dimension impaire N ayant un champ d'hyperplans tangents non dégénéré (de dimension $(N - 1)$ paire). Nous dirons que ces plans sont des *plans de contact.* Tout plan de contact est tangent en un point de la variété de contact. Ce point sera un *point de contact.*

Définition. On appellera *forme de contact* toute forme linéaire définie sur l'espace tangent en un point de la variété de contact, telle que l'ensemble de ses zéros soit un plan de contact.

On remarquera qu'une forme de contact n'est pas une forme différentielle mais une forme algébrique linéaire définie sur un espace tangent.

Définition. On appelle *symplectisée* d'une variété de contact l'ensemble de toutes les formes de contact définies sur une variété de contact, muni de la structure de variété symplectique qui sera définie plus bas.

Remarquons tout d'abord que *l'ensemble de toutes les formes de contact sur une variété de contact possède une structure naturelle de*

variété différentiable de dimension paire $(N + 1)$. Plus exactement, nous pouvons considérer l'ensemble de toutes les formes de contact comme un espace du fibré au-dessus de la variété de contact. La projection sur la base est une application associant à toute forme de contact un point de contact.

Une fibre de ce fibré est un ensemble de formes de contact ayant un point de contact commun. Toutes ces formes se déduisent l'une de l'autre par la multiplication par un nombre non nul (puisqu'elles définissent un même plan de contact). Donc la fibre de notre fibré est unidimensionnelle: c'est une droite sans un point.

Remarquons par ailleurs que sur la variété de toutes les formes de contact opère le groupe multiplicatif des nombres réels non nuls. Plus exactement, le produit d'une forme de contact par un nombre non nul est de nouveau une forme de contact. Le groupe opère sur notre fibré en laissant toute fibre en place (le point de contact reste fixe lorsque la forme est multipliée par un nombre).

R e m a r q u e. Jusqu'ici nous n'avons nulle part utilisé la non-dégénérescence du champ de plans. La non-dégénérescence n'est nécessaire que pour que la variété obtenue par symplectisation soit symplectique.

E x e m p l e. Considérons la variété (de dimension $(2n - 1)$) de tous les éléments de contact d'une variété différentiable de dimension n. Sur la variété des éléments de contact il existe un champ

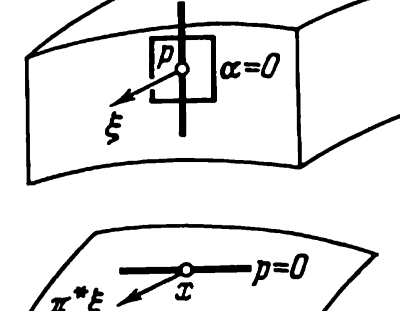



Fig. 237. Symplectisation d'une variété de contact.

d'hyperplans (que nous avons définis plus haut et appelés hyperplans de contact). Donc on peut symplectiser la variété des éléments de contact.

La symplectisation nous donnera une variété de dimension $2n$. Cette variété est un espace du fibré cotangent de la variété initiale de dimension n sans vecteurs nuls. L'action du groupe multiplicatif des nombres réels sur une fibre se ramène à la multiplication des vecteurs de l'espace cotangent par des nombres.

Le fibré cotangent possède une remarquable 1-forme: la forme « $p\, dq$ ». Il existe une 1-forme analogue sur toute variété déduite d'une variété de contact par symplectisation.

1-forme canonique sur la symplectisée d'un espace.

Définition. On appelle 1-*forme canonique* sur la symplectisée d'une variété de contact la 1-forme différentielle α dont la valeur sur tout vecteur ξ tangent à la symplectisée en un point quelconque p (fig. 237) est définie de la manière suivante : considérons p comme une 1-forme sur l'espace tangent à la variété de contact et soit π la projection de la symplectisée sur la variété. Alors

$$\alpha(\xi) = p(\pi_*\xi).$$

Théorème. *La dérivée extérieure de la 1-forme canonique α sur la symplectisée de la variété de contact est une 2-forme non dégénérée.*

Corollaire. *La symplectisée d'une variété de contact possède une structure symplectique qui est canoniquement (i.e. est univoquement et sans aucun arbitraire) définie par la structure de contact de la variété initiale de dimension impaire.*

Démonstration. Cette assertion du théorème étant locale, il suffit de la démontrer dans un petit voisinage d'un point de la variété. Dans un petit voisinage d'un point de la variété de contact on peut définir un champ de plans de contact par la donnée d'une forme différentielle ω sur la variété de contact. Fixons une telle forme ω.

Ce faisant nous avons représenté la partie de la symplectisée située au-dessus du voisinage considéré de la variété de contact sous forme d'un produit cartésien de ce voisinage par une droite sans point.

Plus exactement, au couple (x, λ), où x est un point de la variété de contact et λ un nombre non nul, nous associons une forme de contact qui est définie par la 1-forme différentielle $\lambda\omega$ dans l'espace tangent au point x.

Donc, dans la région considérée de la symplectisée est définie une fonction λ à valeurs non nulles. Soulignons que λ n'est qu'une coordonnée locale sur la variété symplectique et qu'elle n'a pas été définie canoniquement : elle dépend du choix de la 1-forme différentielle ω.

La 1-forme canonique α s'écrit dans les notations adoptées

$$\alpha = \lambda\pi^*\omega$$

et ne dépend pas du choix de ω.

La dérivée extérieure de la 1-forme α est donc

$$d\alpha = d\lambda \wedge \pi^*\omega + \lambda\pi^*\,d\omega.$$

Montrons que la 2-forme $d\alpha$ est non dégénérée, i.e. pour tout vecteur ξ tangent à la symplectisée il existe un vecteur η tel que $d\alpha(\xi, \eta) \neq 0$.

Distinguons parmi les vecteurs tangents à la symplectisée ceux des types suivants. Nous dirons qu'un vecteur ξ est *vertical* s'il est tangent à sa fibre, i.e. si $\pi_*\xi = 0$; le vecteur ξ sera dit *horizontal* s'il est tangent à une surface de niveau de la fonction λ, i.e. si $d\lambda(\xi) = 0$; le vecteur ξ sera appelé *vecteur de contact* si sa projection sur la variété de contact est située dans un plan de contact, i.e. si $\omega(\pi_*, \xi) = 0$ (en d'autres termes si $\alpha(\xi) = 0$).

Calculons la valeur de la forme $d\alpha$ sur le couple de vecteurs ξ, η :

$$d\alpha(\xi, \eta) = (d\lambda \wedge \pi^*\omega)(\xi, \eta) + (\lambda\pi^*\,d\omega)(\xi, \eta).$$

Supposons que le vecteur ξ n'est pas de contact. Prenons pour η un vecteur vertical non nul de sorte que $\pi_*\eta = 0$. Le second terme est nul et le premier égal à

$$-d\lambda(\eta)\,\omega(\pi_*\xi)$$

et non nul, puisque η est un vecteur vertical non nul et ξ n'est pas de contact. Si donc ξ n'est pas un vecteur de contact, alors nous avons trouvé un η tel que $d\alpha(\xi, \eta) \neq 0$.

Supposons que le vecteur ξ est de contact et non vertical. Prenons pour η un vecteur de contact quelconque. Maintenant le premier terme est nul et le second (soit la somme tout entière) se réduit à $\lambda\, d\omega(\pi_*\xi, \pi_*\eta)$. Comme le vecteur ξ n'est pas vertical, le vecteur $\pi_*\xi$ qui est situé dans un plan de contact est non nul. Or la 2-forme $d\omega$ est non dégénérée sur un plan de contact (par définition de la structure de contact). Donc il existe un vecteur de contact η tel que $d\omega(\pi_*\xi, \pi_*\eta) \neq 0$. Comme $\lambda \neq 0$, nous avons trouvé un vecteur η tel que $d\alpha(\xi, \eta) \neq 0$.

Si enfin le vecteur ξ est vertical et non nul, on peut prendre en qualité de η un vecteur quelconque qui ne soit pas de contact. Le théorème est démontré.

R e m a r q u e. La construction des 1-forme α et 2-forme $d\alpha$ vaut pour une variété quelconque ayant un champ d'hyperplans et ne dépend pas des conditions de non-dégénérescence. Cependant la *2-forme $d\alpha$ définira une structure symplectique si seulement le champ de plans est non dégénéré.*

D é m o n s t r a t i o n. Supposons que le champ est dégénéré, i.e. qu'il existe un vecteur non nul ξ' dans un plan du champ tel que $d\omega(\xi', \eta') = 0$ pour tous les vecteurs η' de ce plan. Pour un tel ξ' la quantité $d\omega(\xi', \eta')$ comme fonction de η' est une forme linéaire identiquement nulle sur ce plan. Il existe donc un nombre μ ne dépendant pas de η' tel que

$$d\omega(\xi', \eta') = \mu\omega(\eta')$$

pour tous les vecteurs η' de l'espace tangent au point d'application de ξ'.

Prenons maintenant pour ξ un vecteur tangent à la symplectisée tel que $\pi_*\xi = \xi'$. Un tel vecteur ξ est défini à l'addition près d'un vecteur vertical et nous allons montrer que, si ce vecteur est dûment choisi, on aura

$$d\alpha(\xi, \eta) = 0 \quad \text{pour tous les } \eta.$$

En effet, le premier terme de la formule pour $d\alpha$ est égal à $d\lambda(\xi)\,\omega(\pi_*\eta)$ (puisque $\omega(\pi_*\xi) = 0$). Le second terme est égal à $\lambda d\omega(\pi_*\xi, \pi_*\eta) = \lambda\mu\omega(\pi_*\eta)$. Choisissons la composante verticale du vecteur ξ telle que $d\lambda(\xi) = -\lambda\mu$. Alors le vecteur ξ sera orthogonal gauche à tous les vecteurs η.

Si donc $d\alpha$ est une structure symplectique, alors le champ initial d'hyperplans est une structure de contact et l'assertion mentionnée plus haut est démontrée.

Corollaire. *Le champ d'hyperplans de contact définit une structure de contact sur la variété de tous les élements de contact de toute variété différentiable.*

D é m o n s t r a t i o n. La symplectisée de la variété de dimension $(2n - 1)$ de tous les éléments de contact sur une variété différentiable de dimension n, construite d'après le champ des $(2n - 2)$-plans de contact, est par construction l'espace du fibré cotangent de la variété initiale de dimension n, privé des vecteurs non cotangents. La 1-forme canonique α sur la symplectisée est, en vertu de sa définition, cette 1-forme sur le fibré cotangent que nous avons appelée « $p\, dq$ » et qui repose à la base de la mécanique hamiltonienne (voir § 37). Sa dérivée $d\alpha$ est par conséquent la forme « $dp \wedge dq$ »

qui définit une structure symplectique ordinaire sur l'espace des phases. Donc la forme $d\alpha$ n'est pas dégénérée. Et en vertu de la remarque précédente le champ d'hyperplans de contact n'est pas dégénéré. Ce qui démontre le corollaire.

F. Difféomorphismes de contact et champs de vecteurs.

Définition. On dit qu'un difféomorphisme d'une variété de contact sur elle-même est *de contact* s'il respecte la structure de contact, i.e. s'il transforme tout plan du champ d'hyperplans définissant la structure dans un plan du même champ.

Exemple. Soit la variété de dimension $(2n - 1)$ des éléments de contact d'une variété différentiable de dimension n munie d'une structure de contact ordinaire. A tout élément de contact on peut assigner un « côté positif » en choisissant l'une des deux moitiés en lesquelles cet élément divise l'espace tangent à la variété de dimension n.

Un tel élément de contact sera appelé *élément de contact orienté* (*transversalement*).

Les éléments de contact orientés de notre variété de dimension n forment une variété différentiable de dimension $(2n - 1)$ munie d'une structure naturelle de contact (revêtement double de la variété des éléments de contact ordinaires non orientés).

Supposons maintenant que sur la variété initiale de dimension n est donnée une métrique riemannienne. Alors sur la variété des éléments de contact orientés est défini le « flot géodésique » *). Définissons la transformation dans le temps t de ce flot. A partir du point de contact d'un élément de contact traçons une géodésique perpendiculaire à cet élément et dirigée du côté orientant cet élément. Pendant un intervalle de temps t nous déplacerons le point de contact le long de la géodésique tout en préservant l'orthogonalité de l'élément et de la géodésique. Au bout du temps t nous obtenons un nouvel élément orienté. Ce que nous avons défini est le *flot géodésique des éléments de contact orientés.*

Théorème. *Le flot géodésique des éléments de contact orientés est constitué de difféomorphismes de contact.*

Nous ne démontrerons pas ce théorème, puisque ce n'est pas autre chose que le principe de Huygens formulé dans le nouveau langage (voir § 46).

Définition. On dit qu'un champ de vecteurs défini sur une variété de contact est *de contact* s'il est champ de vitesses d'un groupe (local) à un paramètre de difféomorphismes de contact.

---

*) En toute rigueur il faudrait exiger que la variété riemannienne soit complète, i. e. que les géodésiques soient indéfiniment prolongeables.

Théorème. *Le crochet de Poisson des champs de vecteurs de contact est un champ de vecteurs de contact. Les champs de vecteurs de contact forment une sous-algèbre de l'algèbre de Lie de tous les champs de vecteurs différentiables sur une variété de contact.*

Les démonstrations découlent immédiatement des définitions.

F. Symplectisation des difféomorphismes et champs de contact. D'après chaque difféomorphisme de contact d'une variété de contact on peut construire canoniquement un difféomorphisme symplectique de la symplectisée de cette variété.

Ce difféomorphisme symplectique commute avec l'action du groupe multiplicatif des nombres réels sur la symplectisée et se définit de la manière suivante.

On rappelle qu'un point de la symplectisée d'une variété est une forme de contact sur la variété de contact initiale.

D é f i n i t i o n. On appelle *image d'une forme de contact* p de point de contact x par le difféomorphisme de contact f d'une variété de contact sur elle-même la forme

$$f_! p = (f^*_{f(x)})^{-1} p.$$

En fait tout revient à transporter la forme p de l'espace tangent au point x dans l'espace tangent au point $f(x)$ par le difféomorphisme f (dont la dérivée en x établit un isomorphisme entre ces deux espaces tangents).

La forme $f_! p$ est de contact, car le difféomorphisme f est de contact.

Théorème. *L'application $f_!$ de la symplectisée d'une variété de contact sur elle-même est un difféomorphisme symplectique commutant avec l'action du groupe multiplicatif des nombres réels et respectant la 1-forme canonique α sur la symplectisée.*

D é m o n s t r a t i o n. L'assertion du théorème découle du fait que la 1-forme canonique α, la 2-forme symplectique $d\alpha$ et l'effet du groupe des réels sont définis par la structure de contact même (en les construisant nous n'avons pas utilisé les coordonnées ou autres objets non invariants) ; quant au difféomorphisme f, il respecte la structure de contact. De là il découle que $f_!$ transforme dans lui-même tout ce qui a été intrinsèquement construit d'après la structure de contact et en particulier la 1-forme α, sa dérivée $d\alpha$, l'action du groupe, etc.

Théorème. *Tout difféomorphisme symplectique de la symplectisée d'une variété de contact commutant avec l'action du groupe multiplicatif des réels*: 1) *se projette sur la variété de contact initiale sous forme d'un difféomorphisme de contact*; 2) *respecte la 1-forme canonique α.*

D é m o n s t r a t i o n. Tout difféomorphisme commutant avec l'action du groupe multiplicatif des réels se projette sur un difféomorphisme de la variété de contact. Pour prouver que ce difféomorphisme est de contact il nous suffit de

montrer seulement la deuxième assertion du théorème (puisque sur un plan de contact ne se projettent que les vecteurs ξ tels que $\alpha(\xi) = 0$).

Pour démontrer la deuxième assertion exprimons l'intégrale de la forme α sur un chemin quelconque γ en fonction de la structure symplectique $d\alpha$:

$$\int_{\gamma} \alpha = \lim_{\varepsilon \to 0} \int\int_{\sigma(\varepsilon)} d\alpha,$$

où la 2-chaîne $\sigma(\varepsilon)$ est le produit de γ par tous les nombres possibles de l'intervalle $[\varepsilon, 1]$. Le bord de la 2-chaîne σ est composé, outre γ, de deux segments verticaux et du chemin $\varepsilon\gamma$. Les intégrales de la forme α sur les segments verticaux sont nulles; quant à l'intégrale sur $\varepsilon\gamma$, elle tend vers zéro avec ε.

De l'invariance de la 2-forme $d\alpha$ et de la commutation du difféomorphisme F avec la multiplication par des nombres il résulte que quel que soit le chemin γ

$$\int_{F\gamma} \alpha = \int_{\gamma} \alpha,$$

donc le difféomorphisme F respecte la 1-forme α, c.q.f.d.

Définition. On appelle *symplectisation d'un champ de vecteurs de contact* la construction suivante. Regardons ce champ comme un champ de vitesses d'un groupe à un paramètre de difféomorphismes de contact. Symplectisons les difféomorphismes. On obtient un groupe à un paramètre de difféomorphismes symplectiques. Considérons le champ de vitesses de ce groupe. C'est ce champ qu'on appelle symplectisé du champ de contact initial.

Théorème. *Le symplectisé d'un champ de vecteurs de contact est un champ de vecteurs hamiltonien. Le hamiltonien peut être choisi homogène du premier degré en l'action du groupe multiplicatif des réels:*

$$H(\lambda x) = \lambda H(x).$$

Inversement, tout champ de vecteurs hamiltonien défini sur la symplectisée d'une variété de contact ayant un hamiltonien homogène du premier degré se projette sur la variété de contact initiale sous forme d'un champ de vecteurs de contact.

Démonstration. Le symplectisé du champ de contact est hamiltonien, puisque les symplectisés des difféomorphismes de contact sont symplectiques. L'homogénéité des accroissements du hamiltonien découle de celle des symplectisés des difféomorphismes (de la commutation avec le produit par λ). Donc la première assertion du théorème résulte du théorème de la symplectisation des difféomorphismes de contact.

La seconde partie résulte d'une façon analogue du théorème sur les difféomorphismes symplectiques homogènes, c.q.f.d.

Corollaire. *La symplectisation des champs de vecteurs est une application isomorphe de l'algèbre de Lie des champs de vecteurs de contact*

sur l'algèbre de Lie de tous les champs de vecteurs localement hamiltoniens ayant des hamiltoniens homogènes de degré un.

La démonstration est évidente.

G. Théorème de Darboux pour les structures de contact. Le théorème de Darboux est le théorème d'unicité locale de la structure de contact. On peut l'énoncer sous l'une des trois formes suivantes.

Théorème. *Toutes les variétés de contact de même dimension se correspondent localement par un difféomorphisme de contact* (i.e. il existe un difféomorphisme d'un voisinage suffisamment petit d'un point quelconque d'une variété de contact sur un voisinage d'un point quelconque d'une autre associant au point du premier voisinage le point du second et au champ de plans du premier voisinage le champ de plans du second).

Théorème. *Toute variété de contact de dimension* $(2m - 1)$ *s'applique localement par un difféomorphisme de contact sur la variété des éléments de contact d'un espace de dimension* m.

Théorème. *Toute* 1*-forme différentielle définissant un champ d'hyperplans non dégénéré sur une variété de dimension* $(2n + 1)$ *s'écrit dans un système de coordonnées locales sous la « forme normale »*

$$\omega = x\, dy + dz,$$

où $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ *et* z *sont les coordonnées locales.*

Il est clair que les deux premiers théorèmes découlent du troisième, lequel se déduit du théorème analogue de Darboux sur la forme normale des 2-formes définissant des structures symplectiques (voir § 43).

D é m o n s t r a t i o n d u t h é o r è m e d e D a r b o u x. Symplectisons notre variété. Sur la variété symplectique de dimension $(2n + 2)$ obtenue sont définies la 1-forme canonique α, la 2-forme non dégénérée $d\alpha$, l'application projective π sur la variété de contact initiale et la direction verticale en chaque point.

La 1-forme différentielle ω donnée sur la variété de contact définit une structure de contact en chaque point. Toutes ces formes de contact constituent une sous-variété de dimension $(2n + 1)$ de la variété symplectique. L'application projective π est un difféomorphisme de cette sous-variété sur la variété de contact initiale; quant aux verticales, elles coupent cette sous-variété sous un angle non nul.

Sur la surface construite dans la variété symplectique, considérons le point qui est situé au-dessus du point de la variété de contact qui nous intéresse. Sur la variété symplectique on peut choisir au voisinage de ce point un système de coordonnées local tel que

$$d\alpha = dp_0 \wedge dq_0 + \ldots + dp_n \wedge dq_n$$

et que la surface arithmétique $p_0 = 0$ coïncide avec notre variété de dimension $(2n + 1)$ (voir § 43 où la première coordonnée est arbitrairement choisie dans la démonstration du théorème symplectique de Darboux).

Remarquons maintenant que la 1-forme $p_0\,dq_0 + \ldots + p_n\,dq_n$ admet $d\alpha$ pour dérivée. Donc localement

$$\alpha = p_0\,dq_0 + \ldots + p_n\,dq_n + dw,$$

où w est une fonction que nous pouvons supposer nulle à l'origine des coordonnées. En particulier, sur la surface $p_0 = 0$ la forme α s'écrit

$$\alpha\,|_{p_0=0} = p_1\,dq_1 + \ldots + p_n\,dq_n + dw.$$

L'application projective π permet de transporter les coordonnées $p_1, \ldots, p_n; q_0; q_1, \ldots, q_n$ et la fonction w sur la variété de contact. Plus précisément, nous définissons les fonctions x, y, z au moyen des formules

$$x_i\,(\pi A) = p_i\,(A),\; y_i\,(\pi A) = q_i\,(A), \quad z\,(\pi A) = w\,(A),$$

où A est un point de la surface $p_0 = 0$.

Nous obtenons alors

$$\omega = x\,dy + dz,$$

et il nous reste à vérifier que les fonctions $(x_1, \ldots, x_n; y_1, \ldots, y_n; z)$ forment un système de coordonnées. Pour cela il suffit de vérifier qu'est non nulle la dérivée partielle de la fonction w par rapport à q_0. En d'autres termes, il suffit que la 1-forme α soit non nulle sur le vecteur directeur de l'axe de coordonnées q_0. Cela revient à montrer que la 2-forme $d\alpha$ est non nulle sur un couple de vecteurs dont l'un est porté par l'axe de coordonnées q_0 et l'autre est vertical.

Or le vecteur porté par l'axe q_0 est orthogonal gauche à tous les vecteurs du plan arithmétique $p_0 = 0$. Si de surcroît il était orthogonal gauche au vecteur vertical, il serait orthogonal gauche à tous les vecteurs, en dépit de la non-dégénérescence de la forme $d\alpha$. Donc $\partial w / \partial q_0 \neq 0$ et le théorème est démontré.

H. Hamiltoniens de contact. Supposons que la structure de contact d'une variété de contact est définie par une 1-forme ω et que cette forme est fixée.

D é f i n i t i o n. On appelle *ω-plongement* d'une variété de contact dans sa symplectisée une application qui à tout point de la variété de contact associe la restriction de la forme ω au plan tangent en ce point.

D é f i n i t i o n. On appelle *hamiltonien de contact* d'un champ de vecteurs de contact sur une variété de contact, munie d'une 1-forme ω fixée, une fonction K définie sur la variété de contact dont la valeur en chaque point est égale à celle du hamiltonien homogène H du symplectisé du champ en l'image du point envisagé par le ω-plongement:

$$K\,(A) = H\,(\omega\,|_A).$$

Théorème. *Le hamiltonien de contact K d'un champ de vecteurs X défini sur une variété de contact munie d'une 1-forme ω fixée est égal à la valeur de la forme ω sur ce champ de contact*:

$$K = \omega\,(X).$$

D é m o n s t r a t i o n. Nous allons nous servir de l'expression donnant l'accroissement d'un hamiltonien ordinaire le long d'un chemin en fonction du champ de vecteurs et de la structure de contact (§ 48, C).

A cet effet, à partir du point B de la symplectisée en lequel nous voulons calculer le hamiltonien traçons un segment vertical $\{\lambda B\}$, $0 < \lambda \leqslant 1$. Les déplacements de ce segment pendant un intervalle de temps τ petit par le symplectisé du flot de notre champ X sont partout denses dans une bande $\sigma(\tau)$ de dimension deux. Au point B le hamiltonien prend une valeur égale à la limite

$$H(B) = \lim_{\tau \to 0} \frac{1}{\tau} \int\int_{\sigma(\tau)} d\alpha,$$

puisque $H(\lambda B) \to 0$ lorsque $\lambda \to 0$. Or l'intégrale de la forme $d\alpha$ étendue à la bande est égale à l'intégrale de la 1-forme α étendue au bord engendré par la trajectoire du point B (les autres parties du bord donnent des intégrales nulles). Donc l'intégrale double est tout simplement égale à l'intégrale de la 1-forme α étendue à une portion de trajectoire, et la limite est égale à la valeur de la 1-forme α sur le vecteur vitesse Y du symplectisé du champ. Finalement $K(\pi B) = H(B) = \alpha(Y) = \omega(X)$, c.q.f.d.

I. Formules de calcul. Supposons maintenant que nous utilisons les coordonnées du théorème de Darboux dans lesquelles ω prend la forme normale

$$\omega = x\,dy + dz, \quad x = (x_1, \ldots, x_n), \quad y = (y_1, \ldots, y_n).$$

E x e r c i c e. Trouver les composantes d'un champ de contact dont est donné le hamiltonien de contact $K = K(x, y, z)$.

Réponse. Les équations du champ de contact s'écrivent

$$\begin{cases} \dot{x} = -K_y + xK_z, \\ \dot{y} = K_x, \\ \dot{z} = K - xK_x. \end{cases}$$

S o l u t i o n. Un point du symplectisé peut être défini par $2n + 2$ nombres x_i, y_i, z, λ, où (x, y, z) sont les coordonnées du point de la variété de contact et λ le nombre par lequel il faut multiplier ω pour obtenir le point considéré du symplectisé.

Dans ces coordonnées $\alpha = \lambda x\,dy + \lambda\,dz$. Donc dans le système de coordonnées $\boldsymbol{p}$, $\boldsymbol{q}$, où

$$\boldsymbol{p} = (p, p_0),\ p = \lambda x, \quad p_0 = \lambda,$$
$$\boldsymbol{q} = (q, q_0), \quad q = y, \quad q_0 = z,$$

a forme α prend la forme standard suivante:

$$\alpha = \boldsymbol{p}\,d\boldsymbol{q}, \quad d\alpha = d\boldsymbol{p} \wedge d\boldsymbol{q}.$$

L'action T_μ du groupe multiplicatif des réels s'est ramenée à la multiplication de $\boldsymbol{p}$ par un nombre:

$$T_\mu(\boldsymbol{p}, \boldsymbol{q}) = (\mu\boldsymbol{p}, \boldsymbol{q}).$$

Le hamiltonien de contact K s'exprime en fonction du hamiltonien ordinaire $H = H(p, q, p_0, q_0)$ au moyen de la formule

$$K(x, y, z) = H(x, y, 1, z).$$

La fonction H est homogène du premier degré en p. Donc les dérivées partielles de K au point (x, y, z) sont liées avec les dérivées de H au point $(p = x, p_0 = 1, q = y, q_0 = z)$ par les relations

$$H_q = K_y, \quad H_{q_0} = K_z,$$
$$H_p = K_x, \quad H_{p_0} = K - xK_x.$$

Les équations de Hamilton de hamiltonien H s'écrivent donc en ce point

$$\dot{x} + x\dot{\lambda} = -K_y, \quad \dot{\lambda} = -K_z,$$
$$\dot{y} \quad = \quad K_x, \quad \dot{z} = K - xK_x,$$

d'où l'on déduit la réponse annoncée.

Exercice. Trouver le hamiltonien de contact du crochet de Poisson des champs de vecteurs de contact de hamiltoniens de contact K et K'.

Réponse. $(K, K') + K_z\mathrm{E}K' - K'_z\mathrm{E}K$, où la parenthèse désigne le crochet de Poisson par rapport aux variables x et y, E l'opérateur d'Euler, $\mathrm{E}F = F - xF_x$.

Solution. Il faut exprimer le crochet ordinaire de Poisson des hamiltoniens homogènes H et H' au point $(p = x, p_0 = 1, q = y, q_0 = z)$ en fonction des hamiltoniens de contact K et K' dans les notations de l'exercice précédent. On a

$$(H, H') = H_qH'_p - H_pH'_q = H_qH'_p - H_pH'_q + H_{q_0}H'_{p_0} - H_{p_0}H'_{q_0}.$$

En portant les valeurs des dérivées de la solution de l'exercice précédent on trouve au point considéré

$$(H, H') = K_yK'_x - K_xK'_y + K_z(K' - xK'_x) - K'_z(K - xK_x).$$

J. Variétés de Legendre. Aux variétés lagrangiennes d'un espace des phases symplectique correspond dans la théorie des variétés de contact une classe intéressante de variétés que nous pouvons appeler *legendriennes* puisqu'elles sont intimement liées à la transformation de Legendre.

Définition. On appelle *sous-variété legendrienne* d'une variété de contact de dimension $(2n + 1)$ une variété intégrale de dimension n d'un champ de plans de contact.

En d'autres termes, ce sont les variétés intégrales de plus grande dimension que puisse posséder un champ de plans non dégénéré.

Exemple 1. L'ensemble de tous les éléments de contact tangents à une sous-variété de dimension quelconque d'une variété de dimension m est une sous-variété legendrienne de dimension $(m - 1)$ de la variété de contact de dimension $(2m - 1)$ de tous les éléments de contact.

Exemple 2. L'ensemble de tous les plans tangents au graphe d'une fonction $f = \varphi(x)$ définie sur un espace euclidien de dimension $(n + 1)$ muni des coordonnées $(x_1, \ldots, x_n; f)$ est une sous-variété

legendrienne de l'espace de dimension $(2n+1)$ de tous les hyperplans non verticaux de l'espace du graphe (la structure de contact est définie par la 1-forme

$$\omega = p_1\,dx_1 + \ldots + p_n\,dx_n - df,$$

le plan muni des coordonnées (p, x, f) passe par le point de coordonnées (x, f) parallèlement au plan $f = p_1x_1 + \ldots + p_nx_n$).

La transformation de Legendre se décrit de la manière suivante dans ce langage.

Considérons un autre espace de contact de dimension $(2n+1)$ muni des coordonnées (P, X, F) et d'une structure de contact définie par la forme

$$\Omega = P\,dx - dF.$$

On appelle *involution de Legendre* une application qui au point de coordonnées (p, x, f) du premier espace associe le point

$$P = x,\ X = p,\quad F = px - f$$

du second espace.

On vérifie sans peine que l'involution de Legendre transforme la première structure de contact dans la seconde. Est évident le

Théorème. *Le difféomorphisme d'une variété de contact sur une autre, qui fait se correspondre des plans de contact, fera se correspondre des variétés legendriennes.*

En particulier, l'involution de Legendre transforme une sous-variété legendrienne de plans tangents au graphe de la fonction φ dans une nouvelle variété legendrienne que l'on appelle *transformée de Legendre de la variété initiale.*

La projection de la nouvelle variété sur l'espace muni des coordonnées X, F (parallèlement à la P-direction) n'est pas, en général, une variété différentiable, elle présente des singularités. Cette projection s'appelle *transformée de Legendre du graphe de la fonction* φ.

Si la fonction φ est convexe, alors la projection est elle-même graphe d'une fonction $F = \Phi(X)$. Dans ce cas la fonction Φ est appelée *transformée de Legendre de la fonction* φ.

Comme deuxième exemple considérons le mouvement d'éléments de contact orientés par le flot géodésique sur une variété riemannienne. Prenons pour « front d'onde initial » une sous-variété différentiable quelconque de notre variété riemannienne (la dimension de la sous-variété peut être quelconque). Tous les éléments de contact orientés qui sont tangents à cette sous-variété forment une variété legendrienne dans l'espace de tous les éléments de contact. Du théorème précédent nous déduisons le

Corollaire. *La famille de tous les éléments tangents au front d'onde se transforme par le flot géodésique dans le temps t de nouveau en une variété legendrienne de l'espace de tous les éléments de contact.*

A remarquer que cette nouvelle variété legendrienne n'est pas forcément la famille de tous les éléments tangents à une variété différentiable quelconque, puisque le front d'onde peut présenter des singularités.

Les *singularités legendriennes* peuvent être décrites comme l'ont été les lagrangiennes (voir Appendice 12). Un *fibré legendrien* dans une variété de contact de dimension $(2n + 1)$ est un fibré dont toutes les fibres sont des variétés legendriennes de dimension n. Les singularités legendriennes sont les singularités de l'application projective des sous-variétés legendriennes de dimension n d'une variété de contact de dimension $(2n + 1)$ sur la base de dimension $(n + 1)$ d'un fibré legendrien.

Considérons l'espace $\mathbf{R}^{2n+1}$ muni d'une structure de contact définie par la forme $\alpha = x\,dy + dz$, où $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. L'application projective $(x, y, z) \mapsto (y, z)$ définit un fibré legendrien.

On appelle *équivalence* de fibrés legendriens un difféomorphisme des espaces fibrés, envoyant la structure de contact et les fibres de l'un dans la structure de contact et les fibres de l'autre. On démontre que *tout fibré legendrien est équivalent au fibré spécial que nous venons de décrire* au voisinage de chaque point.

Toute structure de contact d'un espace fibré définit sur les fibres une structure locale d'espace projectif. Les équivalences legendriennes respectent cette structure, i.e. définissent des transformations localement projectives de fibres.

Le théorème qui suit permet de décrire localement les sous-variétés et les applications legendriennes à l'aide des fonctions génératrices.

Théorème. *Pour toute partition $I + J$ de l'ensemble des indices $(1, \ldots, n)$ en deux sous-ensembles disjoints et pour toute fonction $S(x_I, y_J)$ de n variables x_i, $i \in I$ et y_j, $j \in J$, les formules*

$$y_I = \frac{\partial S}{\partial x_I}, \quad x_J = -\frac{\partial S}{\partial y_J}, \quad z = S - x_I \frac{\partial S}{\partial x_I}$$

définissent une sous-variété legendrienne dans $\mathbf{R}^{2n+1}$. Inversement, toute sous-variété legendrienne de $\mathbf{R}^{2n+1}$ est définie au voisinage de chacun de ses points par les formules précédentes au moins pour l'un des 2^n choix possibles de la sous-variété I.

La démonstration se base sur le fait que sur une variété legendrienne $dz + x\,dy = 0$, donc $d(z + x_I y_I) = y_I\,dx_I - x_J dy_J$.

Remplaçons dans les formules du théorème précédent S par les fonctions de la liste des singularités lagrangiennes élémentaires insérée dans l'Appendice 12. On obtient des singularités legendriennes qui se conservent sous l'effet de petites déformations de l'application legendrienne $(x, y, z) \mapsto (y, z)$ (i.e. se transforment en des

singularités équivalentes sous l'effet d'une faible déformation de la fonction S). Pour $n < 6$ toute application legendrienne se transforme par un petit déplacement en une application dont toutes les singularités sont localement équivalentes aux singularités de la liste A_k $(1 \leqslant k \leqslant 6)$, D_k $(4 \leqslant k \leqslant 6)$, E_6.

En particulier, nous obtenons une liste de singularités des fronts d'onde génériques dans des espaces de dimension inférieure à 7.

Dans l'espace ordinaire cette liste est

$$A_1: S = \pm x_1^2, \quad A_2: S = \pm x_1^3; \quad A_3: S = \pm x_1^4 + x_1^2 y_2,$$

où $I = \{1\}$, $J = \{2\}$, $n = 2$.

Les projections de ces variétés legendriennes sur la base du fibré legendrien (i.e. sur l'espace décrit par les coordonnées y_1, y_2, z) possèdent respectivement un point simple dans le cas de A_1, une arête de rebroussement dans le cas A_2 et une forme de queue d'aronde (voir fig. 246) dans le cas A_3.

Donc, le front d'onde générique possède dans l'espace tridimensionnel seulement des arêtes de rebroussement et des points de type « queue d'aronde ». Lorsque le front est en mouvement, on observe à certains moments des bifurcations des trois types A_4, D_4^-, D_4^+ (cf. Appendice 12 où sont représentées les caustiques correspondantes balayées par les singularités du front lorsque celui-ci est en mouvement).

Exercice 1. Portons un segment de longueur t sur chaque normale intérieure à une ellipse dans le plan. Tracer la courbe obtenue et étudier ses singularités et leurs bifurcations lorsque t varie.

Exercice 2. Même exercice pour un ellipsoïde à trois axes dans l'espace ordinaire.

K. Contactisation. Outre la simplectisation des variétés de contact il existe la contactisation des variétés symplectiques possédant une structure symplectique homologue à zéro.

La contactisée E^{2n+1} d'une variété symplectique (M^{2n}, ω^2) se construit comme un espace fibré de fibre $\mathbf{R}$ au-dessus de M^{2n}.

Soit U un voisinage suffisamment petit d'un point x de M sur lequel on a défini un système de coordonnées canoniques p, q, de sorte que $\omega = dp \wedge dq$. Considérons le produit cartésien $U \times \mathbf{R}$ muni des coordonnées p, q, z.

Soit $V \times \mathbf{R}$ le produit cartésien construit pour un autre (ou pour le même) voisinage V sur lequel on a défini des coordonnées P, Q, Z; $\omega = dP \wedge dQ$. Si les voisinages U et V sur M se coupent, nous identifions les fibres au-dessus des points d'intersection dans les deux produits de telle sorte que la forme $dz + p\,dq = dZ + P\,dQ = \alpha$ soit définie globalement (ceci est possible puisque $P\,dQ - p\,dq$ est une différentielle totale sur $U \cap V$).

On vérifie aisément que par collage on obtient le fibré E^{2n+1} au-dessus de M^{2n} et que la forme α définit une structure de contact sur E. La variété E s'appelle *contactisée* de la variété symplectique M. Si la classe d'homologies de la forme ω^2 est entière, on peut définir une contactisée de fibre S^1.

L. Intégration des équations aux dérivées partielles du premier ordre. Soit M^{2n+1} une variété de contact, E^{2n} une hypersurface dans M^{2n+1}. La structure de contact de M définit une structure géométrique sur E, en particulier un champ de directions caractéristiques. L'analyse de cette structure géométrique permet de ramener l'intégration des équations non linéaires générales aux dérivées partielles du premier ordre à l'intégration d'un système d'équations différentielles ordinaires.

Nous supposons que la variété E^{2n} est transversale en tous ses points aux plans de contact. Alors l'intersection du plan tangent à E^{2n} en chaque point avec le plan de contact est de dimension $(2n - 1)$, de sorte que sur E^{2n} est défini un champ d'hyperplans. De plus la structure de contact sur M^{2n+1} définit sur E^{2n} un champ de droites contenues dans les $(2n - 1)$-plans mentionnés.

En effet, soient α une 1-forme sur M^{2n+1} définissant localement une structure de contact; $\omega = d\alpha$; $\mathbf{R}^{2n}$ le plan de contact au point $x \in E^{2n}$ et enfin $\Phi = 0$ une équation locale sur E^{2n} (de sorte que $d\Phi$ n'est pas nulle en x).

La restriction de $d\Phi$ à $\mathbf{R}^{2n}$ définit une forme linéaire non nulle sur $\mathbf{R}^{2n}$. La 2-forme ω définit sur $\mathbf{R}^{2n}$ une structure d'espace vectoriel symplectique et partant un isomorphisme entre cet espace et son dual. Donc à la 1-forme $d\Phi \,|\, \mathbf{R}^{2n}$ non nulle correspond un vecteur ξ non nul de $\mathbf{R}^{2n}$, de sorte que $d\Phi(\cdot) = \omega(\xi, \cdot)$. Le vecteur ξ est appelé *vecteur caractéristique* de la variété E^{2n} au point x. Le vecteur caractéristique ξ est contenu dans l'intersection de $\mathbf{R}^{2n}$ avec le plan tangent à E^{2n} de sorte que $d\Phi(\xi) = 0$.

Le vecteur ξ n'est pas défini de façon unique par la variété E^{2n} et la structure de contact sur M, mais à la multiplication près par un nombre non nul. En effet, la 2-forme ω sur $\mathbf{R}^{2n}$ comme la 1-forme $d\Phi$ sur $\mathbf{R}^{2n}$ sont définies à la multiplication près par un nombre non nul.

Les directions des vecteurs caractéristiques sont définies de façon unique par la structure de contact en chaque point de la variété E^{2n}.

Donc un champ de directions caractéristiques est défini sur l'hypersurface E^{2n} de la variété de contact M. Les courbes intégrales de ce champ de directions sont appelées *caractéristiques*.

Soit maintenant donnée une sous-variété I de dimension $(n - 1)$ de l'hypersurface E^{2n}, qui est intégrale pour le champ de contact (de sorte que le plan tangent à I en chacun de ses points appartient au plan de contact).

Théorème. *Si en un point x de I une caractéristique sur E^{2n} n'est pas tangente à I, alors au voisinage de ce point les caractéristiques passant par les points de I forment une sous-variété legendrienne L^n dans M^{2n+1}.*

D é m o n s t r a t i o n. Soit ξ le champ des vecteurs caractéristiques sur E^{2n}. En vertu de la formule d'homotopie (cf. page 168) on a sur E^{2n}

$$L_\xi \alpha = d i_\xi \alpha + i_\xi d\alpha.$$

Or $i_\xi \alpha = 0$, puisque le vecteur caractéristique appartient au plan de contact. Donc sur E^{2n} on a $L_\xi \alpha = i_\xi \omega$. Mais la 1-forme $i_\xi \omega$ est nulle sur l'intersection du plan tangent à E^{2n} avec le plan de contact (car sur le plan de contact $i_\xi \omega = d\Phi$ et sur le plan tangent $d\Phi = 0$). Donc sur le plan tangent à E^{2n} on a $i_\xi \omega = c\alpha$. Par conséquent sur l'hypersurface E^{2n}

$$L_\xi \alpha = c\alpha$$

(où c est une fonction différentiable au voisinage du point x).

Soit maintenant $\{g^t\}$ le flot (local) du champ ξ et soit η un vecteur tangent à E^{2n}. Posons $\eta(t) = g^t_* \eta$ et $y(t) = \alpha(\eta(t))$. Alors la fonction y obéit à l'équation différentielle linéaire

$$dy/dt = c(t)\, y(t).$$

Si $\eta(0)$ est tangent à I, alors $y(0) = \alpha(\eta(0)) = 0$. Cela veut dire que $y(t) = \alpha(\eta(t)) = 0$, i.e. $\eta(t)$ est contenu dans le plan de contact pour tous les t. Donc $g^t I$ est une variété intégrale du champ de contact. Par conséquent, la variété engendrée par tous les $\{g^t I\}$ pour des petits t est legendrienne. Le théorème est démontré.

E x e m p l e. Soit l'espace $\mathbf{R}^{2n+1}$ muni des coordonnées $x_1, \ldots, x_n; p_1, \ldots, p_n; u$ et d'une structure de contact définie par la 1-forme $\alpha = du - p\,dx$. La fonction $\Phi(x, p, u)$ définit une équation différentielle $\Phi(x, \partial u/\partial x, u) = 0$ et la sous-variété $E = \Phi^{-1}(0)$ dans l'espace $\mathbf{R}^{2n+1}$ (appelé *espace de* 1-*fibres* de fonctions dans $\mathbf{R}^n$).

On appelle *condition initiale* de l'équation $\Phi = 0$ la donnée des valeurs f de la fonction u sur une hypersurface Γ de dimension $(n - 1)$ dans l'espace de dimension n muni des coordonnées $x_1, \ldots, x_n$.

La condition initiale définit les dérivées de u suivant $(n - 1)$ directions indépendantes en chaque point de Γ. La dérivée suivant une direction transversale à Γ peut être d'une façon générale calculée à partir de l'équation ; si de plus sont réalisées les conditions du théorème des fonctions implicites, alors la condition initiale est dite *non caractéristique*.

Une condition initiale non caractéristique définit une sous-variété intégrale I de dimension $(n - 1)$ munie de la forme α (qui est le graphe de l'application $u = f(x)$, $p = p(x)$, $x \in \Gamma$). Les caractéristiques sur E qui coupent I forment une sous-variété legendrienne dans $\mathbf{R}^{2n+1}$, qui est le graphe de l'application $u = u(x)$, $p = \partial u/\partial x$. La fonction $u(x)$ obtenue est solution de l'équation $\Phi(x, \partial u/\partial x, u) = 0$ avec la condition initiale $u|_\Gamma = f$.

Remarquons que pour trouver la fonction u il faut simplement résoudre un système de $2n$ équations différentielles du premier ordre pour les caractéristiques sur E et effectuer une série d'opérations « algébriques ».

APPENDICE 5

SYSTÈMES DYNAMIQUES À SYMÉTRIE

En vertu du théorème de E. Noether les groupes à un paramètre de symétries d'un système dynamique définissent les intégrales premières. Si le système se conserve par un groupe plus large de symétries, il s'introduit alors plusieurs intégrales.

Les variétés de niveau communes de ces intégrales premières sont des variétés invariantes du flot. Tout sous-groupe d'un groupe de symétries laissant fixe cette variété invariante opère sur elle. Dans nombre de cas on peut étudier la variété quotient de la variété invariante par ce sous-groupe. Cette variété quotient est appelée espace des phases réduit. L'espace des phases réduit possède une structure symplectique naturelle. Le système dynamique hamiltonien initial définit de nouveau sur lui un système hamiltonien.

La partition de l'espace des phases en variétés de niveau communes des intégrales premières présente en général des singularités. Exemple : la partition du plan de phase en lignes de niveau d'énergie.

Dans cet appendice on se propose d'étudier succinctement des systèmes dynamiques dans les espaces des phases réduits et leurs liens avec les variétés invariantes de l'espace des phases primitif. Toutes ces questions ont déjà été traitées par Jacobi et Poincaré (« exclusion du nœud » dans les problèmes de plusieurs corps, « abaissement de l'ordre » dans les systèmes à symétrie, « rotations permanentes » du solide, etc.). On trouvera un exposé plus détaillé dans le langage actuel dans les articles : S. Smale, *Topology and mechanics*, « Inventiones Mathematicae » **10** : 4 (1970), 305-331, **11** : 1 (1970) 45-64, et I. Marsden, A. Weinstein, *Reduction of symplectic manifolds with symetries.*

A. Actions poissonniennes des groupes de Lie. Considérons une variété symplectique (M^{2n}, ω^2) et soit G un groupe de Lie opérant sur cette variété comme un groupe de difféomorphismes symplectiques. Tout sous-groupe à un paramètre du groupe G opère alors comme un flot localement hamiltonien sur M. Dans de nombreux cas importants ces flots possèdent des hamiltoniens univoques.

E x e m p l e. Soient V une variété différentiable, G un groupe quelconque de Lie de difféomorphismes de V. Tout difféomorphisme envoie une 1-forme sur V dans une 1-forme. Donc le groupe G opère sur le fibré cotangent $M = T^*V$.

On rappelle que sur un fibré cotangent il existe toujours une 1-forme canonique α (la forme $p\,dq$) et une structure symplectique naturelle $\omega = d\alpha$. Le groupe G opère symplectiquement sur M puisqu'il respecte la 1-forme α et par conséquent la 2-forme $d\alpha$.

Un sous-groupe à un paramètre $\{g^t\}$ du groupe G définit un flot sur M. Il est immédiat de vérifier que ce flot possède un hamiltonien univoque. Plus exactement, le hamiltonien H est donné par la formule du théorème de Noether

$$H(x) = \alpha\left(\frac{d}{dt}\Big|_{t=0} g^t x\right), \quad \text{où } x \in M.$$

Supposons maintenant qu'est donné un groupe de Lie G opérant symplectiquement sur la variété symplectique connexe M de telle sorte qu'à tout élément a de l'algèbre de Lie du groupe G il associe un groupe à un paramètre de difféomorphismes symplectiques d'hamiltonien H_a univoque. Ces hamiltoniens sont définis à l'addition près de termes constants qui peuvent être choisis de façon à faire dépendre H_a linéairement de a. Pour cela il suffit de choisir arbitrairement les constantes dans les hamiltoniens pour des vecteurs de base quelconques de l'algèbre de Lie du groupe G et ensuite de définir le hamiltonien pour tout élément de l'algèbre comme une combinaison linéaire de vecteurs de base.

Donc d'après l'action symplectique du groupe de Lie G ayant des hamiltoniens univoques sur M on peut construire une application linéaire de l'algèbre de Lie de G dans l'algèbre de Lie des hamiltoniens sur M. Ceci étant, au commutateur de deux éléments de l'algèbre de Lie est associée une fonction $H_{[a,\, b]}$ égale au crochet de Poisson (H_a, H_b) *ou n'en différant que par l'addition d'une constante*:

$$H_{[a,\, b]} = (H_a, H_b) + C(a, b).$$

R e m a r q u e. La présence de la constante C dans cette formule est due à un phénomène intéressant: l'existence d'une classe bidimensionnelle de cohomologies de l'algèbre de Lie des champs (globalement) hamiltoniens.

La quantité $C(a, b)$ est une fonction antisymétrique bilinéaire sur l'algèbre de Lie. L'identité de Jacobi entraîne

$$C([a, b], c) + C([b, c], a) + C([c, a], b) = 0.$$

Une fonction bilinéaire antisymétrique sur l'algèbre de Lie jouissant de cette propriété s'appelle *cocycle de dimension deux* de l'algèbre de Lie.

Si l'on choisit d'autres constantes dans les fonctions de Hamilton, au lieu du cocycle C, on obtiendra le cocycle C' où

$$C'(a, b) = C(a, b) + p([a, b])$$

et où p est une fonction linéaire sur l'algèbre de Lie.

Un tel cocycle C' est appelé cocycle *cohomologue* au cocycle C.

La classe des cocycles cohomologues est appelée *classe de cohomologies* de l'algèbre de Lie.

Donc, l'*action symplectique du groupe G telle qu'il existe des hamiltoniens univoques définit une classe de cohomologies de dimension deux de l'algèbre de Lie*

du groupe G. Cette classe de cohomologies mesure les écarts par rapport à l'action pour laquelle le hamiltonien du commutateur peut être choisi égal au crochet de Poisson des hamiltoniens.

Définition. L'action d'un groupe connexe de Lie sur une variété symplectique est dit *poissonnienne* si les hamiltoniens des groupes à un paramètre sont univoques et choisis tels que le hamiltonien dépende linéairement d'un élément de l'algèbre de Lie et le hamiltonien du commutateur soit égal au crochet de Poisson des hamiltoniens:

$$H_{[a,\,b]} = (H_a,\ H_b).$$

En d'autres termes, l'action poissonnienne d'un groupe définit un homomorphisme de l'algèbre de Lie de ce groupe dans l'algèbre de Lie des hamiltoniens.

Exemple. Soit V une variété différentiable, G un groupe de Lie quelconque opérant sur V comme un groupe de difféomorphismes. Soit $M = T^*V$ le fibré cotangent de la variété V munie de la structure symplectique ordinaire $\omega = d\alpha$. Définissons les hamiltoniens des groupes à un paramètre comme précédemment:

$$H_a(x)=\alpha\left(\frac{d}{dt}\bigg|_{t=0} g^t x\right), \quad x\in T^*V. \tag{1}$$

Théorème. *L'action construite est poissonnienne.*

Démonstration. Par définition de la 1-forme α les hamiltoniens H_a sont linéaires et homogènes en les « impulsions » (i.e. sur chaque espace cotangent). Donc leurs crochets de Poisson sont linéaires et homogènes. Et la fonction

$$H_{[a,\,b]} - (H_a,\ H_b)$$

est linéaire et homogène en les impulsions. Etant constante, elle est nulle, c.q.f.d.

On vérifie de façon analogue que la symplectisée de toute action de contact est une action poissonnienne.

Exemple. Soient V l'espace euclidien ordinaire, G le groupe de dimension six des mouvements de cet espace. Une base de l'algèbre de Lie est engendrée par les six groupes à un paramètre: les translations de vitesse unité le long des axes de coordonnées q_1, q_2, q_3 et les rotations de vitesse angulaire unité autour de ces axes. Les hamiltoniens correspondants sont, en vertu de la formule (1) (dans les notations usuelles), égaux à p_1, p_2, p_3; M_1, M_2, M_3, où $M_1 = q_2p_3 - q_3p_2$, etc. Le théorème démontré signifie que les crochets de Poisson de ces six fonctions prises deux à deux sont égaux aux hamiltoniens des commutateurs des groupes à un paramètre correspondants.

L'action poissonnienne du groupe G sur la variété symplectique M définit une application de la variété M dans l'espace dual de l'algèbre de Lie du groupe

$$P: M \to \mathfrak{g}^*.$$

Plus exactement, fixons un point x de M et considérons sur l'algèbre de Lie une fonction qui à tout élément a de l'algèbre de Lie associe la valeur du hamiltonien H_a en x:

$$p_x(a) = H_a(x).$$

Cette fonction linéaire p_x sur l'algèbre de Lie est précisément cet élément de l'algèbre $\mathfrak{g}^*$ qui est associé au point x:

$$P(x) = p_x.$$

J. M. Souriau a proposé d'appeler l'application P *moment*. Remarquons que la valeur du moment est toujours un vecteur de l'espace vectoriel $\mathfrak{g}^*$.

Exemple. Soient V une variété différentiable, G un groupe de Lie opérant sur V comme un groupe de difféomorphismes, $M = T^*V$ le fibré cotangent, H_a le hamiltonien de l'action poissonnienne de G sur M construit plus haut (voir (1)).

Alors le moment $P: M \to \mathfrak{g}^*$ peut être décrit de la manière suivante. Soit l'application $\Phi: G \to M$ définie en faisant opérer tous les éléments du groupe G sur un point fixe x de M (de sorte que $\Phi(g) = gx$). La 1-forme canonique α sur M induit sur G la 1-forme $\Phi^*\alpha$. Sa restriction à l'espace tangent à G en l'unité est une forme linéaire sur l'algèbre de Lie.

Donc à tout point x de M on a associé une forme linéaire sur l'algèbre de Lie. Il est immédiat de vérifier que l'application obtenue est le moment de l'action poissonnienne considérée.

En particulier, si V est l'espace euclidien ordinaire et G le groupe des rotations autour de O, alors les valeurs de ce moment sont les vecteurs ordinaires du moment cinétique; si G est le groupe des rotations autour d'un axe, alors les valeurs du moment sont les moments cinétiques par rapport à cet axe; si enfin G est le groupe des déplacements parallèles, alors les valeurs du moment sont les vecteurs impulsion.

Théorème. *L'action poissonnienne d'un groupe de Lie connexe G se transforme par le moment P en l'action coadjointe du groupe G sur l'algèbre* $\mathfrak{g}^*$ (cf. Appendice 2), *i. e. le diagramme*

$$\begin{array}{ccc} M & \xrightarrow{g} & M \\ {\scriptstyle P}\downarrow & & \downarrow{\scriptstyle P} \\ \mathfrak{g}^* & \xrightarrow{\mathrm{Ad}_g^*} & \mathfrak{g}^* \end{array}$$

est commutatif.

Corollaire. *Supposons que le hamiltonien $H: M \to \mathbf{R}$ est invariant par l'action poissonnienne de groupe G sur M. Alors le moment est intégrale première d'un système de hamiltonien H.*

Démonstration du théorème. Le théorème affirme que le hamiltonien H_a du groupe à un paramètre h^t se transforme par le difféomorphisme g en le hamiltonien $H_{\mathrm{Ad}_g a}$ du groupe à un paramètre gh^tg^{-1}. Soit g^s un groupe à un paramètre de hamiltonien H_b. Il suffit de démontrer que les dérivées par rapport à s (pour $s = 0$) des fonctions $H_a(g^sx)$ et $H_{\mathrm{Ad}_{g^s}a}(x)$ coïncident. La première dérivée est égale à la valeur du crochet de Poisson des fonctions (H_a, H_b) au point x. La deuxième est $H_{[a,b]}(x)$. Comme l'effet est poissonnien, le théorème est démontré.

Démonstration du corollaire. La dérivée de toute composante du moment suivant la direction du flot de hamiltonien H est nulle,

puisqu'elle est égale à la dérivée de la fonction H suivant la direction du flot du sous-groupe à un paramètre correspondant du groupe G, c.q.f.d.

B. Espace des phases réduit. Soit donnée une action poissonnienne du groupe G sur une variété symplectique M. Considérons un ensemble de niveau du moment, i.e. la contre-image d'un point quelconque $p \in \mathfrak{g}^*$ par l'application P. Notons cet ensemble M_p de sorte que (fig. 238)

$$M_p = P^{-1}(p).$$

Dans nombre de cas importants, l'ensemble M_p est une variété. Ce sera le cas si p est une valeur régulière du moment, i.e. si la différentielle de l'application P en chaque point de l'ensemble M_p applique tout espace tangent à M dans tout l'espace tangent à $\mathfrak{g}^*$.

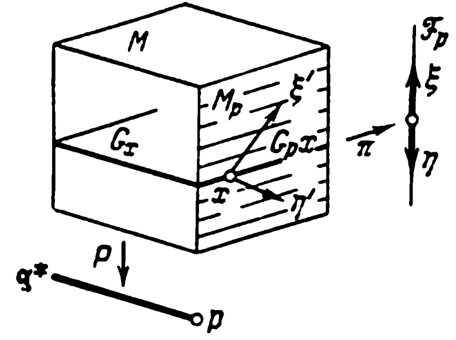


Fig. 238. Espace des phases réduit.

Le groupe de Lie G qui opère sur M en général permute les ensembles M_p. Cependant le sous-groupe stationnaire du point p de la représentation coadjointe (i. e. le sous-groupe qui est composé de tous les éléments g du groupe G tels que $\mathrm{Ad}^*_g p = p$) laisse M_p en place. Désignons ce sous-groupe stationnaire par G_p. Le groupe G_p est un groupe de Lie qui évolue sur l'ensemble de niveau du moment M_p.

L'espace des phases réduit se déduit de M_p : c'est l'espace des orbites du groupe G_p qui opère sur M_p. Pour que cet espace soit correctement défini il faut que soient réalisées certaines conditions. Par exemple, il suffit de supposer que

1) p est une valeur régulière de sorte que M_p est une variété;
2) le sous-groupe stationnaire G_p est compact;
3) les éléments du groupe G_p opèrent sur M_p sans laisser de points fixes.

R e m a r q u e. Ces conditions peuvent être allégées. Par exemple, au lieu de la compacité du groupe G_p on peut exiger que l'action soit propre (i.e. que les contre-images des ensembles compacts par l'application $(g, x) \to (g(x), x)$ soient compactes). Ainsi le groupe opère toujours proprement sur lui-même par des translations à gauche ou à droite.

Si les conditions 1), 2), 3) sont réalisées, on définit sans peine une structure de variété différentiable sur l'ensemble des orbites

de l'action de G_p sur M_p. Plus exactement, toute surface transversale à l'orbite $G_p x$ et de dimension égale à la codimension de l'orbite peut servir de carte au voisinage d'un point $x \in M_p$.

La variété d'orbites obtenue s'appelle *espace des phases réduit d'un système à symétrie.*

Désignons par F_p l'espace des phases réduit correspondant à la valeur p du moment. La variété F_p est la base du fibré $\pi: M_p \to F_p$ de fibre difféomorphe au groupe G_p.

L'espace des phases réduit F_p est muni d'une structure symplectique naturelle. Plus exactement, considérons deux vecteurs quelconques ξ, η tangents à F_p en f. Le point f est une orbite du groupe G_p sur la variété M_p. Soit x un point de cette orbite. Les vecteurs ξ et η tangents à l'orbite se déduisent de vecteurs ξ', η' tangents en x à M_p par l'application projective $\pi: M_p \to F_p$.

Définition. On appelle *produit scalaire gauche* des vecteurs ξ, η tangents à l'espace des phases réduit en un de ses points le produit scalaire gauche des vecteurs correspondants ξ', η' tangents à la variété symplectique initiale M:

$$[\xi, \eta]_p = [\xi', \eta'].$$

Théorème *). *Le produit scalaire gauche des vecteurs ξ, η ne dépend pas du choix du point x et des représentants ξ', η' et définit une structure symplectique sur l'espace des phases réduit.*

Corollaire. *L'espace des phases réduit est de dimension paire.*

Démonstration du théorème. Considérons dans l'espace tangent en x à M les deux espaces suivants:

$T(M_p)$ — l'espace tangent à la variété de niveau du moment M_p;
$T(Gx)$ — l'espace tangent à l'orbite Gx du groupe G.

Lemme. *Ces deux espaces sont complémentaires orthogonaux gauches l'un de l'autre relativement à TM.*

Démonstration. Un vecteur ζ est contenu dans le complémentaire orthogonal gauche du plan tangent à une orbite du groupe G si et seulement si les produits scalaires gauches du vecteur ζ et des vecteurs vitesses des flots hamiltoniens du groupe G sont nuls (par définition). Or ces produits scalaires gauches sont égaux aux dérivées des fonctions de Hamilton correspondantes suivant la direction de ζ. Donc le vecteur ζ est contenu dans l'orthocomplément gauche d'une orbite du groupe G si et seulement si la dérivée du moment suivant la direction de ζ est égale à zéro, i.e. si ζ est contenu dans $T(M_p)$. La première assertion du lemme est démontrée; la deuxième est évidente.

Les représentants ξ' et η' sont définis à l'addition près d'un vecteur du plan tangent à l'orbite du groupe G_p qui passe par les points d'application de ξ' et η'. Or ce plan tangent est l'intersection des plans tangents à l'orbite Gx et à la

---

*) Ce théorème a été énoncé pour la première fois sous cette forme par Marsden et Weinstein. De nombreux cas particuliers ont été étudiés depuis les temps de Jacobi et utilisés par Poincaré en mécanique, par Kirillov et Costant dans la théorie des groupes et par Faddéev dans la relativité générale.

variété M_p (en vertu du dernier théorème du point A). Donc l'addition à ξ' d'un vecteur de $T(G_p x)$ ne modifie pas les produits scalaires gauches par tous les vecteurs η' de $T(M_p)$ (puisqu'en vertu du lemme, $T(G_p x)$ est orthogonal gauche à $T(M_p)$). Donc l'indépendance par rapport au choix des représentants ξ', η' est démontrée.

L'indépendance de la quantité $[\xi, \eta]_p$ par rapport au choix du point x sur l'orbite f découle de la symplecticité de l'action du groupe G sur M et de l'invariance de M_p. Donc sur F_p est définie la 2-forme différentielle

$$\Omega_p(\xi, \eta) = [\xi, \eta]_p.$$

Elle n'est pas dégénérée. En effet, si $[\xi, \eta]_p = 0$ pour tous les η, alors le représentant correspondant ξ' est orthogonal gauche à tous les vecteurs de $T(M_p)$. Donc ξ' appartient à l'orthocomplément gauche de $T(M_p)$ par rapport à TM. Et en vertu du lemme $\xi' \in T(Gx)$, i.e. $\xi = 0$.

La forme Ω_p est fermée. Pour s'en assurer considérons une carte quelconque, i.e. une surface de M_p coupant transversalement l'orbite du groupe G_p en un point.

La forme Ω_p est représentée sur cette carte par une 2-forme induite par la 2-forme ω, qui définit la structure symplectique sur l'espace M tout entier, par plongement de cette surface. Comme la forme ω est fermée, la forme induite l'est également. Le théorème est démontré.

E x e m p l e 1. Soit $M = \mathbf{R}^{2n}$ un espace euclidien de dimension $2n$ muni des coordonnées p_k, q_k et de la 2-forme $\sum dp_k \wedge dq_k$. Soit $G = S^1$ un cercle; l'action de G sur M est donnée par le hamiltonien de l'oscillateur linéaire

$$H = \frac{1}{2} \sum (p_k^2 + q_k^2).$$

Alors le moment est tout simplement $H: \mathbf{R}^{2n} \to \mathbf{R}$, la variété de niveau nul du moment est une sphère S^{2n-1} et l'espace quotient, l'espace projectif complexe CP^{n-1}.

Donc le théorème précédent définit une structure symplectique sur l'espace projectif complexe. On vérifie sans peine que cette structure coïncide (à un facteur près) avec celle que nous avons construite dans l'Appendice 3.

E x e m p l e 2. Soit V un groupe de Lie, G le même groupe mais opérant par des translations à gauche. Alors M_p est une sous-variété du fibré tangent au groupe formée par les vecteurs qui par translation à droite dans l'unité du groupe donnent un même élément dans l'algèbre $\mathfrak{g}^*$.

Donc la variété M_p est difféomorphe au groupe lui-même et est une section invariante à droite du fibré cotangent. Toutes les valeurs p sont régulières.

Le sous-groupe stationnaire G_p du point p est composé des éléments du groupe pour lesquels les translations à gauche et à droite de l'élément p donnent le même résultat. Les actions des éléments distincts de l'unité du groupe G_p sur M_p ne possèdent pas de points fixes puisqu'ils n'existent pas pour les translations à droite du groupe dans lui-même).

Le groupe G_p opère proprement (voir remarque de la page 379). Donc l'espace des orbites du groupe G_p sur M_p est une variété symplectique.

Mais cet *espace d'orbites* s'identifie aisément à *l'orbite du point p dans la représentation coadjointe*. En effet, appliquons la section invariante *à droite* M_p du fibré cotangent dans l'espace cotangent au groupe en l'unité par translations *à gauche*. On obtient l'application

$$\pi : M_p \to \mathfrak{g}^*.$$

L'image de cette application est l'orbite du point p dans la représentation coadjointe, et les fibres, les orbites de l'action du groupe G_p. La structure symplectique de l'espace des phases réduit définit donc une structure symplectique sur les orbites de la représentation coadjointe.

On vérifie par un calcul immédiat que c'est la structure traitée dans l'Appendice 3.

E x e m p l e 3. Supposons que le groupe G est un cercle S^1 et qu'il opère sans points fixes sur la variété V. Alors est définie une action poissonnienne du cercle sur le fibré cotangent $M = T^*V$. Nous pouvons définir les variétés de niveau du moment M_p (de codimension 1 dans M) et les variétés quotients F_p (dont les dimensions sont de deux unités inférieures à celle de M).

En outre nous pouvons construire l'espace quotient de l'espace de configuration V en identifiant l'un à l'autre tous les points de chaque orbite du groupe sur V. Désignons par W cet espace quotient.

Théorème. *L'espace des phases réduit F_p est symplectiquement difféomorphe au fibré cotangent de l'espace de configuration quotient W.*

D é m o n s t r a t i o n. Soit π: $V \to W$ la factorisation, $\omega \in T^*W$ une 1-forme sur W au point $w = \pi v$. La forme $\pi^*\omega$ sur V au point v appartient à M_0 et après la factorisation définit un point dans F_0. Inversement, les éléments de F_0 sont des 1-formes invariantes sur V, nulles sur les orbites; elles définissent des 1-formes sur W. Donc, nous avons construit une application $T^*W \to F_0$; on voit aisément que c'est un difféomorphisme symplectique.

On ramène le cas $p \neq 0$ au cas $p = 0$ en procédant comme suit. Considérons sur V une métrique riemannienne invariante par G. L'intersection de M_p avec le plan cotangent à V en v est un hyperplan. La forme quadratique qui définit la métrique possède sur cet hyperplan un unique point de minimum : $s(v)$. La soustraction du vecteur $s(v)$ transforme l'hyperplan $M_p \cap T^*V_v$ dans $M_0 \cap T^*V_v$ et nous obtenons le difféomorphisme $F_p \to F_0$. Le théorème est démontré.

C. Applications à l'étude des rotations stationnaires et des bifurcations des variétés invariantes. Soit donnée une action poissonnienne du groupe G sur une variété symplectique M et soit H une fonction sur M invariante par G. Soit F_p l'espace des phases réduit (nous supposons que les conditions qui le définissent sont réalisées).

Le champ hamiltonien de fonction de Hamilton H est tangent à chaque variété de niveau du moment M_p (puisque ce moment est

intégrale première). Ce champ qui est invariant par G_p définit sur l'espace des phases réduit F_p un champ que l'on appelle *champ réduit.*

Théorème. *Le champ réduit sur l'espace des phases réduit est hamiltonien. La fonction de Hamilton du champ réduit prend en tout point de l'espace des phases réduit une valeur qui est égale à celle de la fonction de Hamilton initiale au point correspondant de l'espace des phases initial.*

Démonstration. La relation qui définit le champ hamiltonien X_H de hamiltonien H sur la variété M munie de la forme ω, soit

$$dH(\xi) = \omega(X_H, \xi) \quad \text{pour tous les } \xi,$$

entraîne une relation analogue pour le champ réduit en vertu de la définition de la structure symplectique sur F_p, c.q.f.d.

Exemple. Considérons un solide asymétrique ayant un point fixe et soumis à l'action de la pesanteur (ou à une autre force potentielle symétrique par rapport à un axe vertical).

Sur l'espace de configuration SO(3) opère le groupe S^1 des rotations autour de la verticale. La fonction de Hamilton est invariante dans les rotations, donc il s'introduit un système réduit sur l'espace des phases réduit.

L'espace des phases réduit est ici le fibré cotangent à l'espace de configuration quotient (voir exemple 3 de la page 382). La construction de l'espace des orbites de l'espace de configuration sur lequel opère le groupe des rotations autour de la verticale fut réalisée par Poisson de la manière suivante.

Définissons la position du solide par le repère orthonormé (e_1, e_2, e_3). Les trois composantes verticales des vecteurs du repère définissent un vecteur dans l'espace euclidien arithmétique de dimension trois. Le module de ce vecteur est 1 (pourquoi?). Ce *vecteur de Poisson* *) γ définit le repère initial aux rotations près autour de la verticale (pourquoi?).

Donc l'espace de configuration quotient est une sphère S^2 et l'espace des phases réduit, le fibré cotangent à la sphère T^*S^2.

Sur le fibré cotangent à la sphère la fonction de Hamilton réduite est la somme de l'« énergie cinétique du mouvement réduit » (cette énergie est une forme quadratique en les vecteurs cotangents) et d'un « potentiel effectif » (qui comprend l'énergie potentielle et l'énergie cinétique de rotation autour de la verticale).

Le passage à l'espace des phases réduit se ramène dans ce cas à l'« exclusion de la coordonnée cyclique φ » à cette différence près que la procédure ordinaire

*) Poisson a démontré que les équations du mouvement d'un solide pesant s'expriment en fonction du vecteur γ sous la forme remarquablement simple des « équations d'Euler-Poisson »

$$\frac{dM}{dt} - [M, \omega] = \mu g [\gamma, l], \quad \frac{d\gamma}{dt} = [\gamma, \omega].$$

d'exclusion implique que l'espace de configuration ou l'espace des phases soit un produit cartésien par un cercle alors qu'ici nous avons seulement un fibré. Ce fibré devient un produit cartésien si l'on rejette quelques sous-variétés de l'espace de configuration (i.e. si l'on introduit des coordonnées à singularités aux pôles) ; l'avantage de l'approche proposée est qu'il n'existe aucune singularité réelle (excepté celle du système de coordonnées) au voisinage des pôles.

D é f i n i t i o n. Les courbes de phase de M qui se projettent sur les positions d'équilibre du système réduit sur l'espace des phases réduit F_p s'appellent *équilibres relatifs* du système initial.

E x e m p l e. Les rotations stationnaires d'un solide fixé au centre d'inertie sont des équilibres relatifs.

D'une façon analogue sont équilibres relatifs les mouvements d'un solide pesant qui se réduisent à une rotation à vitesse constante autour d'un axe vertical.

Théorème. *Une courbe de phase d'un système de fonction de Hamilton G-invariante est un équilibre relatif si et seulement si elle est orbite d'un sous-groupe à un paramètre du groupe G dans l'espace des phases initial.*

D é m o n s t r a t i o n. Qu'une courbe de phase qui soit orbite se projette en un point est chose évidente. Si une courbe de phase $x(t)$ se projette en un point alors on peut la représenter de façon unique sous la forme $x(t) = g(t)\, x(0)$ et l'on voit aussitôt que $\{g(t)\}$ est un sous-groupe, c.q.f.d.

Corollaire 1. *Un solide asymétrique fixé en un point de l'axe d'un champ potentiel quelconque à symétrie axiale admet au moins deux rotations stationnaires (pour chaque valeur du moment cinétique par rapport à l'axe de symétrie).*

Corollaire 2. *Un solide à symétrie axiale fixé en un point de son axe de symétrie admet dans un champ potentiel quelconque au moins deux rotations stationnaires (pour chaque valeur du moment cinétique par rapport à l'axe de symétrie).*

Les deux corollaires découlent de ce qu'une fonction possède sur la sphère au moins deux points critiques.

On utilise aussi les équilibres relatifs pour l'étude des modifications topologiques des variétés invariantes de l'énergie et du moment.

Théorème. *Les points critiques de l'application du moment-énergie*

$$P \times H : M \to \mathfrak{g}^* \times \mathbf{R}$$

sur un ensemble régulier de niveau du moment sont exactement les équilibres relatifs.

D é m o n s t r a t i o n. Les points critiques de l'application $P \times H$ sont les extremums liés de la fonction H sur une variété de niveau du moment M_p (puisque la variété de niveau considérée est régulière, i.e. quel que soit $x \in M_p$ on a $P_* TM_x = T\mathfrak{g}^*_p$).

Les extremums liés de la fonction H sur M_p définissent sur l'espace des orbites de G_p les points critiques de la fonction réduite de Hamilton (puisque H est invariante par le groupe G_p). Le théorème est démontré.

En fait l'étude des équilibres relatifs et des singularités de l'application énergie-moment n'est pas simple et n'a pas été menée jusqu'au bout même pour le problème classique du mouvement d'un solide asymétrique dans un champ de gravitation. Le cas où le centre de gravité est situé sur un des axes d'inertie a été traité par S. Katoc dans l'appendice proposé à la traduction de l'article de Smale cité plus haut.

Dans ce problème l'espace des phases est de dimension 6, le groupe est un cercle; l'espace des phases réduit T^*S^2 est de dimension 4.

Dans l'espace des phases réduit les variétés de niveau d'énergie régulier (selon les valeurs du moment et de l'énergie) sont de quatre types: S^3, $S^2 \times S^1$, RP^3 et la sphère S^3 à deux anses

$$S^1 \times D^2 \ (D^2 \text{ est un disque, } \{(x, y) : x^2 + y^2 \leqslant 1\}).$$

APPENDICE 6

FORMES NORMALES DES HAMILTONIENS QUADRATIQUES

Dans le présent appendice on trouvera la table des formes normales auxquelles on peut ramener un hamiltonien quadratique par une transformation symplectique réelle. Cette table a été composée par D. Galine sur la base de l'article de J. Williamson, *On algebraic problem, concerning the normal forms of linear dynamical systems*, *Amer. J. of Math.* 58, n° 1 (1936), 141-163.

Williamson indique les formes normales auxquelles on peut ramener une forme quadratique dans un espace symplectique sur un champ quelconque.

A. Notations. Ecrivons le hamiltonien sous la forme

$$H = \frac{1}{2}(Ax, x)$$

où $x = (p_1, \ldots, p_n; q_1, \ldots, q_n)$ est un vecteur écrit dans une base symplectique, A un opérateur linéaire symplectique. Les équations canoniques prennent alors la forme

$$\dot{x} = I\,Ax \text{ où } I = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}.$$

Nous appellerons *valeurs propres* du hamiltonien les valeurs propres de l'opérateur infinitésimal symplectique IA. D'une façon analogue, par bloc de Jordan nous entendrons un bloc de Jordan de l'opérateur IA.

Les valeurs propres du hamiltonien sont de quatre types: les couples réels $(a, -a)$, les couples imaginaires purs $(ib, -ib)$, les quaternes $(\pm a \pm ib)$ et les valeurs propres nulles.

Les blocs de Jordan correspondant aux deux termes d'un couple ou aux quatre termes d'un quaterne possèdent toujours la même structure.

Lorsque la partie réelle d'une valeur propre est nulle il importera de distinguer l'ordre pair et l'ordre impair. Les blocs d'ordre impair à valeur propre nulle sont en nombre pair et on peut les répartir en couples.

Voici la liste des formes normales.

B. Hamiltoniens. A tout couple de blocs de Jordan d'ordre k à valeurs propres réelles $\pm a$ est associé le hamiltonien

$$H = -a\sum_{j=1}^{k} p_j q_j + \sum_{j=1}^{k-1} p_j q_{j+1}.$$

A un quaterne de blocs de Jordan d'ordre k à valeurs propres $\pm a \pm bi$, le hamiltonien

$$H = -a\sum_{j=1}^{2k} p_j q_j + b\sum_{j=1}^{k} (p_{2j-1}q_{2j} - p_{2j}q_{2j-1}) + \sum_{j=1}^{2k-2} p_j q_{j+2}.$$

A un couple de blocs de Jordan d'ordre impair k à valeurs propres nulles, le hamiltonien

$$H = \sum_{j=1}^{k-1} p_j q_{j+1} \quad (k=1,\ H=0).$$

A un bloc de Jordan d'ordre pair $2k$ à valeur propre nulle, l'un des deux hamiltoniens suivants:

$$H = \pm\frac{1}{2}\left(\sum_{j=1}^{k-1} p_j p_{k-j} - \sum_{j=1}^{k} q_j q_{k-j+1}\right) - \sum_{j=1}^{k-1} p_j q_{j+1}.$$

Pour $k = 1$, $H = \pm\frac{1}{2}q_1^2$. Les hamiltoniens de signes différents ne se transforment pas l'un dans l'autre.

A un couple de blocs de Jordan d'ordre impair $(2k+1)$ à valeurs propres imaginaires pures $\pm bi$, l'un des deux hamiltoniens suivants:

$$H = \pm\frac{1}{2}\left[\sum_{j=1}^{k} (b^2 p_{2j}p_{2k-2j+2} + q_{2j}q_{2k-2j+2}) - \right.$$
$$\left. - \sum_{j=1}^{k+1} (b^2 p_{2j-1}p_{2k-2j+3} + q_{2j-1}q_{2k-2j+3})\right] - \sum_{j=1}^{2k} p_j q_{j+1}.$$

Pour $k = 0$ $H = \pm 1/2\,(b^2p_1^2 + q_1^2)$. Les hamiltoniens de signes différents ne se transforment pas l'un dans l'autre.

A un couple de blocs de Jordan d'ordre pair $2k$ à valeurs propres imaginaires pures $\pm bi$, l'un des deux hamiltoniens suivants:

$$H = \pm\frac{1}{2}\left[\sum_{j=1}^{k}\left(\frac{1}{b^2} q_{2j-1}q_{2k-2j+1} + q_{2j}q_{2k-2j+2}\right) - \right.$$
$$\left. - \sum_{j=1}^{k-1} (b^2 p_{2j+1}p_{2k-2j+1} + p_{2j+2}p_{2k-2j+2})\right] - b^2\sum_{j=1}^{k} p_{2j-1}q_{2j} + \sum_{j=1}^{k} p_{2j}q_{2j-1}.$$

Pour $k=1$ $H = \pm 1/2\left(\frac{1}{b^2}q_1^2 + q_2^2\right) - b^2p_1q_2 + p_2q_1.$

Les hamiltoniens de signes différents ne se transforment pas l'un dans l'autre par une application symplectique réelle.

Théorème de Williamson. *L'espace symplectique vectoriel réel sur lequel est définie une forme quadratique H se décompose en une somme directe de sous-espaces symplectiques réels deux à deux orthogonaux gauches telle que la forme H soit la somme des formes définies plus haut sur ces sous-espaces.*

C. Blocs de Jordan inévitables. Un hamiltonien individuel générique ne possède pas de valeurs propres multiples et se ramène à une forme simple (tous les blocs de Jordan sont d'ordre 1). Cependant si l'on considère non pas un hamiltonien individuel, mais toute une famille de hamiltoniens dépendant de paramètres, alors pour certaines valeurs exceptionnelles de ces paramètres il peut s'introduire des structures de Jordan plus compliquées. On peut se débarrasser de certaines d'entre elles par une petite déformation de la famille, les autres sont inévitables et ne font que se déformer légèrement. Si le nombre l de paramètres de la famille est fini, ces cas inévitables sont en nombre fini dans la famille l-paramétrique. Le théorème de Galine énoncé plus bas permet d'énumérer tous ces cas quel que soit l fixe.

Soient $n_1(z) \geqslant n_2(z) \geqslant \ldots \geqslant n_s(z)$ les dimensions des blocs de Jordan à valeur propre $z \neq 0$, et $m_1 \geqslant m_2 \geqslant \ldots \geqslant m_u$ et $\tilde{m}_1 \geqslant \tilde{m}_2 \geqslant \ldots \geqslant \tilde{m}_v$ les dimensions des blocs de Jordan à valeur propre nulle, en outre m_j sont pairs et $\tilde{m}_j$ impairs (dans tout couple de blocs de dimension impaire on ne considère qu'un).

Théorème. *La variété de hamiltoniens dont les blocs de Jordan ont les dimensions indiquées possède dans l'espace de tous les hamiltoniens la codimension*

$$c = \frac{1}{2} \sum_{z \neq 0} \left[\sum_{j=1}^{s(z)} (2j-1)\, n_j(z) - 1 \right] +$$

$$+ \frac{1}{2} \sum_{j=1}^{u} (2j-1)\, m_j + \sum_{j=1}^{v} [2(2j-1)\, \tilde{m}_j + 1] +$$

$$+ 2 \sum_{j=1}^{u} \sum_{k=1}^{v} \min \{m_j, \tilde{m}_k\}.$$

(Remarquons que si zéro n'est pas valeur propre, seul le premier terme de la somme est non nul.)

Corollaire. *Dans les familles de systèmes linéaires hamiltoniens dépendant, d'une façon générale, de l paramètres on ne trouve que des systèmes ayant des blocs de Jordan tels que le nombre c calculé au moyen*

de la formule précédente ne soit pas supérieur à l: tous les cas où $c > l$ s'éliminent par une petite déformation de la famille.

Corollaire. *Dans les familles à un et à deux paramètres les blocs de Jordan inévitables sont nécessairement de l'un des 12 types suivants:*

$$l = 1 : (\pm a)^2, \quad (\pm ia)^2, \ 0^2$$

(les blocs de Jordan sont désignés ici par leurs déterminants; par exemple $(\pm a)^2$ désigne un couple de blocs de Jordan d'ordre 2 et à valeurs propres a et $-a$ respectivement):

$$l = 2 : (\pm a)^3, \ (\pm ai)^3, \ (\pm a \pm bi)^2, \ 0^4, \ (\pm a)^2(\pm b)^2,$$
$$(\pm ai)^2(\pm bi)^2, \ (\pm a)^2(\pm bi)^2, \ (\pm a)^2\, 0^2, \ (\pm ai)^2\, 0^2$$

(les autres valeurs propres sont simples).

Galine a également calculé les formes normales auxquelles on peut ramener toute famille de systèmes hamiltoniens linéaires dépendant différentiablement de paramètres au moyen d'un changement de coordonnées linéaire symplectique dépendant différentiablement de ces paramètres. Par exemple, pour le bloc de Jordan $(\pm a)^2$ une telle forme normale du hamiltonien sera

$$H(\lambda) = -a\,(p_1q_1 + p_2q_2) + p_1q_2 + \lambda_1 p_1 q_1 + \lambda_2 p_2 q_1$$

(λ_1 et λ_2 sont des paramètres).

APPENDICE 7

FORMES NORMALES DE SYSTÈMES HAMILTONIENS AU VOISINAGE DE POINTS FIXES ET DES TRAJECTOIRES FERMÉES

Pour étudier le comportement des solutions de l'équation de Hamilton au voisinage d'une position d'équilibre on ne peut pas toujours se limiter à l'équation linéarisée. En effet, les positions asymptotiquement stables des systèmes hamiltoniens sont impossibles en vertu du théorème de Liouville sur la conservation du volume. Donc la stabilité du système linéarisé est toujours neutre : les valeurs propres de la partie linéaire du hamiltonien d'un champ de vecteurs dans une position d'équilibre stable sont toutes situées sur l'axe imaginaire.

Pour les systèmes génériques d'équations différentielles cette stabilité neutre peut être perturbée par l'addition de termes non linéaires arbitrairement petits. Pour les systèmes hamiltoniens la situation est plus compliquée. Supposons par exemple que la partie quadratique de la fonction de Hamilton (qui définit la partie linéaire du champ de vecteurs) est définie positive ou négative dans une position d'équilibre. Alors la fonction de Hamilton possède un maximum ou un minimum dans cette position. Donc cette position d'équilibre est stable (au sens de Liapounov mais pas asymptotiquement) non seulement pour le système linéarisé mais pour tout le système non linéaire.

Cependant la partie quadratique de la fonction de Hamilton n'est pas forcément définie positive ou négative dans la position d'équilibre stable. Un exemple simple est celui de la fonction $H = p_1^2 + q_1^2 - p_2^2 - q_2^2$. L'étude de la stabilité de systèmes possédant une telle partie quadratique doit tenir compte des termes suivants du développement taylorien de la fonction de Hamilton et avant tout des termes au cube (i. e. des termes au carré des vecteurs du champ de la vitesse). Cette étude est commode à effectuer si la fonction de Hamilton (et par conséquent le champ de vecteurs hamiltonien) est ramenée à la forme la plus simple possible par un changement canonique de variables convenable. En d'autres termes, pour étudier les solutions il faut choisir un système de coordonnées canoniques au voisinage de la position d'équilibre qui permette d'écrire

la fonction de Hamilton et les équations de mouvement dans une forme plus simple.

Ce problème se résout facilement pour les champs de vecteurs non hamiltoniens: ici le cas général est celui où, au voisinage de la position d'équilibre, le champ de vecteurs est linéaire dans un système de coordonnées convenable (les théorèmes de Poincaré et Siegel correspondants figurent par exemple dans l'ouvrage: C. L. Siegel, *Lectures on Celestial Mechanics*, Springer, 1971).

Dans le cas hamiltonien la situation est plus compliquée. La première difficulté est que d'une façon générale il est impossible de ramener un champ hamiltonien à une forme linéaire normale par un changement de coordonnées canonique. Plus exactement, d'ordinaire on peut éliminer les termes au cube de la fonction de Hamilton mais pas tous ceux du quatrième degré (car dans un système linéaire la fréquence des oscillations ne dépend pas de l'amplitude alors qu'elle en dépend en général dans le cas non linéaire). On s'affranchit de cette difficulté en choisissant une forme normale non linéaire qui prenne en considération la variation des fréquences. En définitive (dans le cas de non-résonance) on peut introduire au voisinage de la position d'équilibre des variables action-angle telles que le système soit intégrable aux termes près de puissance aussi grande que l'on veut du développement taylorien.

Ceci permet d'étudier le comportement du système sur de grands intervalles de temps pour des conditions initiales proches de la position d'équilibre. Cependant cela ne suffit pas pour établir si une position d'équilibre est stable au sens de Liapounov (car sur un intervalle de temps infini l'influence du terme résiduel négligé dans la série de Taylor peut perturber la stabilité). Cette stabilité aurait lieu si l'on pouvait réduire le système hamiltonien à une forme normale exacte sans avoir à négliger les termes résiduels. Cependant on démontre que cette réduction exacte est, d'une façon générale, impossible; quant aux séries formelles pour les transformations canoniques réduisant le système à une forme normale, elles sont réellement *divergentes* dans le cas général.

La divergence de ces séries est liée au fait que la réduction à une forme normale entraînerait un comportement plus simple des courbes de phase (elles devraient être des hélices quasi périodiques de tores). Le comportement des courbes de phase au voisinage d'une position d'équilibre d'un système hamiltonien est discuté dans l'Appendice 8. Dans le présent appendice on donne les résultats formels sur la réduction à la forme normale aux termes près de degré supérieur.

L'idée de la réduction des systèmes hamiltoniens aux formes normales remonte à Lindstedt et Poincaré *); les formes normales

*) H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, v. 1, Œuvres choisies. Dover publications inc., New York, 1957.

au voisinage de la position d'équilibre ont été étudiées en détail par Birkhoff (G. D. Birkhoff, *Dynamic Systems*, New York, 1927).

Les formes normales pour les cas dégénérés sont traitées par A. Bruno, *Forme analytique des équations différentielles* (en russe). Travaux de la société mathématique de Moscou, vv. 25 et 26.

A. Forme normale d'un système conservatif au voisinage d'une position d'équilibre. Supposons que dans l'approximation linéaire la position d'équilibre d'un système hamiltonien à n degrés de liberté est stable et que les n fréquences propres $\omega_1, \ldots, \omega_n$ sont toutes distinctes. Alors la partie quadratique du hamiltonien se ramène par une transformation canonique linéaire à la forme

$$H = \frac{1}{2}\omega_1(p_1^2+q_1^2) + \ldots + \frac{1}{2}\omega_n(p_n^2+q_n^2).$$

(Certains ω_k peuvent être négatifs.)

Définition. Les fréquences propres $\omega_1, \ldots, \omega_n$ vérifient une relation de résonance d'ordre K s'il existe des nombres entiers k non tous nuls tels que

$$k_1\omega_1 + \ldots + k_n\omega_n = 0, \quad |k_1| + \ldots + |k_n| = K.$$

Définition. On appelle *forme normale de Birkhoff de degré s pour un hamiltonien* un polynôme de degré s des coordonnées canoniques (P_l, Q_l) qui en réalité est un polynôme (de degré $[s/2]$) des variables $\tau_l = (P_l^2 + Q_l^2)/2$.

Pour un système à un degré de liberté par exemple la forme normale de degré $2m$ (ou $2m+1$) s'écrit

$$H_{2m} = H_{2m+1} = a_1\tau + a_2\tau^2 + \ldots + a_m\tau^m, \quad \tau = (P^2+Q^2)/2,$$

et pour un système à deux degrés de liberté la forme normale de Birkhoff de degré 4 sera

$$H_4 = a_1\tau_1 + a_2\tau_2 + a_{11}\tau_1^2 + 2a_{12}\tau_1\tau_2 + a_{22}\tau_2^2.$$

Les coefficients a_1 et a_2 sont fréquences propres, les coefficients a_{ij} lient les fréquences aux amplitudes.

Théorème. *Supposons que toutes les fréquences propres ω_l ne vérifient aucune relation de résonance d'ordre $\leqslant s$. Il existe alors un système canonique de coordonnées au voisinage de la position d'équilibre tel que la fonction de Hamilton s'y écrive sous la forme de Birkhoff de degré s aux termes près de degré $s+1$:*

$$H(p, q) = H_s(P, Q) + R, \quad R = O(|P| + |Q|)^{s+1}.$$

Ce théorème se démontre facilement dans un système de coordonnées complexes

$$z_l = p_l + iq_l, \quad w_l = p_l - iq_l.$$

(Lorsqu'on passe à ce système il importe de multiplier le hamiltonien par $-2i$.) Si les termes de degré inférieur à N qui ne figurent pas dans la forme normale

sont déjà éliminés, alors un changement de fonction génératrice $P_q + S_N(P, q)$ (où S_N est un polynôme de degré N) n'affectera dans le développement taylorien que les termes de degré $\geqslant N$.

Le coefficient du terme de degré N de la fonction de Hamilton, soit

$$z_1^{\alpha_1} \ldots z_n^{\alpha_n} w_1^{\beta_1} \ldots w_n^{\beta_n} \quad (\alpha_1 + \ldots + \alpha_n + \beta_1 + \ldots + \beta_n = N)$$

devient par ce changement

$$s_{\alpha\beta} [\lambda_1 (\beta_1 - \alpha_1) + \ldots + \lambda_n (\beta_n - \alpha_n)],$$

où $\lambda_l = i\omega_l$ et $s_{\alpha\beta}$ est le coefficient de $z^\alpha w^\beta$ dans le développement de la fonction $S_N(p, q)$ suivant les variables z, w.

Puisque nous avons admis qu'il n'y avait pas résonance, le coefficient entre crochets de $s_{\alpha\beta}$ est non nul, sauf si notre terme de degré N s'exprime en fonction des produits $z_l w_l = 2\tau_l$ (i.e. lorsque tous les α_l sont égaux à β_l). Nous pouvons donc éliminer tous les termes de degré N à l'exception de ceux qui sont fonctions de τ_l. En faisant $N = 3, 4, \ldots, s$ nous obtenons le résultat annoncé.

En se servant du théorème de Birkhoff il est utile de remarquer qu'un système dont le hamiltonien est une forme normale est intégrable. Plus exactement, considérons les « coordonnées canoniques polaires » τ_l, φ_l en fonction desquelles P_l et Q_l s'expriment par les formules

$$P_l = \sqrt{2\tau_l} \cos \varphi_l, \quad Q_l = \sqrt{2\tau_l} \sin \varphi_l.$$

Comme le hamiltonien ne s'exprime qu'en fonction des coordonnées d'action τ_l, le système est intégrable et décrit des mouvements quasi périodiques de fréquences $\omega = \partial H/\partial \tau$ sur les tores $\tau =$ const. En particulier, la position d'équilibre $P = Q = 0$ est stable pour la forme normale.

B. Forme normale d'une transformation canonique au voisinage d'un point fixe. Considérons une application canonique (i.e. qui conserve l'aire) d'un 2-plan sur lui-même. Supposons que cette transformation laisse en place l'origine des coordonnées et que sa partie linéaire admet $\lambda = e^{\pm i\alpha}$ pour valeurs propres (i. e. est une rotation d'angle α dans une base symplectique convenable de coordonnées p, q). Nous dirons que cette transformation est *elliptique*.

Définition. *La forme normale de Birkhoff de degré s pour une transformation* est par définition une transformation canonique d'un plan dans lui-même, rotation d'un angle variable qui est un polynôme de degré non supérieur à $m = [s/2] - 1$ de la variable d'action d'un système de coordonnées polaires canonique :

$$(\tau, \varphi) \longmapsto (\tau, \varphi + \alpha_0 + \alpha_1 \tau + \ldots + \alpha_m \tau^m),$$

où

$$p = \sqrt{2\tau} \cos \varphi, \quad q = \sqrt{2\tau} \sin \varphi.$$

Théorème 2. *Si une valeur propre λ d'une transformation canonique elliptique n'est pas racine de l'unité d'indice $\leqslant s$, alors cette trans-*

formation se ramène par un changement canonique de variables à la forme normale de Birkhoff de degré s aux termes près de degré $\geqslant s+1$.

La généralisation de la transformation elliptique est le produit cartésien de n rotations elliptiques de plans (p_l, q_l) de valeurs propres $\lambda_l = e^{\pm i\alpha_l}$. La forme normale de Birkhoff de degré s est définie par la formule

$$(\tau, \varphi) \mapsto (\tau, \varphi + \partial S/\partial\tau),$$

où S est un polynôme de degré non supérieur à $[s/2]$ des variables d'action $\tau_1, \ldots, \tau_n$.

Théorème 3. *Si les valeurs propres* λ_l *d'une transformation canonique elliptique multidimensionnelle sont libres des résonances*

$$\lambda_1^{k_1} \ldots \lambda_n^{k_n} = 1, \quad |k_1| + \ldots + |k_n| \leqslant s,$$

alors cette transformation se ramène à la forme normale de Birkhoff de degré s (aux termes près de degré s du développement taylorien de cette application au point $p = q = 0$).

C. Formes normales d'une équation à coefficients périodiques au voisinage des positions d'équilibre. Soit $p = q = 0$ la position d'équilibre d'un système ayant une fonction de Hamilton 2π-périodique en le temps. Supposons que l'équation linéarisée a été ramenée par une transformation symplectique linéaire périodique en le temps à la forme normale autonome de fréquences propres $\omega_1, \ldots, \omega_n$.

Nous dirons que le système est *résonnant d'ordre* $K > 0$ s'il existe une relation

$$k_1\omega_1 + \ldots + k_n\omega_n + k_0 = 0$$

avec $k_0, k_1, \ldots, k_n$ entiers telle que $|k_1| + \ldots + |k_n| = K$.

Théorème. *Si le système est libre de résonance d'ordre* $\leqslant s$, *il existe une transformation canonique* 2π*-périodique en le temps qui, au voisinage de la position d'équilibre, ramène ce système à la forme normale de Birkhoff de degré s qui lui correspondrait s'il était autonome à la seule différence que les termes résiduels* R *de degré* $\geqslant s+1$ *dépendront périodiquement du temps.*

Soit enfin donnée une trajectoire fermée d'un système autonome d'équations de Hamilton. Nous pouvons alors ramener ce système au voisinage de cette trajectoire à une forme normale par l'un des deux procédés suivants:

1. Réduction isoénergétique: fixons la constante d'énergie et considérons un voisinage d'une trajectoire fermée d'une variété de dimension $(2n-1)$ de niveau d'énergie comme l'espace des phases élargi d'un système à $(n-1)$ degrés de liberté dépendant périodiquement du temps.

2. Surface d'une section: fixons la constante d'énergie et la valeur de l'une des coordonnées (de sorte que la trajectoire fermée coupe transversalement le $(2n - 2)$-plan obtenu). Alors les courbes de phase voisines de la courbe donnée définissent une application de ce $(2n - 2)$-plan sur lui-même ayant un point fixe sur la trajectoire fermée. Cette application respecte la structure symplectique naturelle sur le $(2n - 2)$-plan et nous pouvons l'étudier à l'aide de la forme normale du point B.

Lorsqu'on étudie les trajectoires fermées des systèmes autonomes hamiltoniens on se trouve en présence d'une situation nouvelle par rapport à l'étude des positions d'équilibre des systèmes à coefficients périodiques. En effet, les trajectoires fermées des systèmes autonomes ne sont pas isolées, mais forment (en général) des familles à un paramètre. Le paramètre d'une famille est la valeur de la constante d'énergie. Supposons en effet que pour une valeur arbitrairement choisie de la constante d'énergie la trajectoire fermée coupe transversalement le $(2n - 2)$-plan construit plus haut. Alors il existera une trajectoire fermée analogue pour les valeurs proches de la constante d'énergie. En vertu du théorème des fonctions implicites nous pouvons même affirmer que cette trajectoire dépend différentiablement de la valeur de la constante d'énergie.

Si maintenant nous voulons nous servir de la forme normale de Birkhoff pour étudier une famille à un paramètre de trajectoires fermées nous nous heurterons à la difficulté suivante. Lorsqu'on modifie le paramètre de la famille les valeurs propres du problème linéarisé varieront en général. Donc pour certaines valeurs du paramètre nous serons inévitablement placés dans des cas de résonances qui interdiront la réduction à la forme normale.

Les résonances d'ordre inférieur sont particulièrement dangereuses puisqu'elles influent sur les premiers termes du développement taylorien. Si nous nous intéressons à la trajectoire fermée pour laquelle les valeurs propres sont proches d'une relation de résonance d'ordre inférieur, alors il faut légèrement modifier la forme normale de Birkhoff. Plus exactement, dans le cas d'une résonance d'ordre N certaines des expressions

$$k_0 - [\omega_1(\beta_1 - \alpha_1) + \ldots \omega_n(\beta_n - \alpha_n)], \quad |\alpha| + |\beta| = N,$$

par lesquelles il faut diviser la fonction de Hamilton pour éliminer les termes de degré N s'annulent. Pour les valeurs du paramètre voisines de la résonance, la combinaison de fréquences propres indiquée n'est en général pas nulle, mais très petite (pour cette raison on l'appelle « petit diviseur »).

La division par un petit diviseur entraîne que

la transformation qui réduit à la forme normale dépend discontinûment du paramètre (elle possède un pôle pour la valeur de résonance du paramètre);

le domaine dans lequel la forme normale de Birkhoff décrit bien le système se contracte en un point pour la valeur de résonance.

Pour s'affranchir de ces insuffisances, il ne faut pas éliminer certains termes du hamiltonien (notamment ceux qui deviennent résonnants pour la valeur de résonance du paramètre). *Il faut les garder non seulement pour la valeur de résonance du paramètre mais aussi pour les valeurs voisines* *).

La forme normale ainsi obtenue est plus compliquée que la forme ordinaire, mais dans nombre de cas elle peut fournir une intéressante information sur le comportement des solutions au voisinage de la résonance.

D. Exemple: étude de la résonance d'ordre 3. Comme exemple simple voyons ce qui se passe avec la trajectoire fermée d'un système hamiltonien autonome à deux degrés de liberté au voisinage d'une valeur de la constante d'énergie telle que la période d'oscillations des trajectoires voisines près de la trajectoire fermée soit le triple de la période de révolution sur la trajectoire fermée.

En vertu de ce qui précède, ce problème se ramène à l'étude d'une famille à un paramètre de systèmes hamiltoniens non autonomes à un degré de liberté, 2π-périodiques en le temps, au voisinage de la position d'équilibre. Cette position d'équilibre peut être placée à l'origine des coordonnées pour toutes les valeurs du paramètre (par un changement de variables dépendant du paramètre).

D'autre part, le système linéarisé à l'équilibre peut être transformé en système linéaire à coefficients constants par un changement de coordonnées canonique linéaire 2π-périodique en le temps. Dans ces coordonnées le flot du système linéarisé est une rotation uniforme autour de la position d'équilibre. La vitesse angulaire ω de cette rotation dépend du paramètre.

Pour la valeur de résonance du paramètre, $\omega = 1/3$ (i. e. un tiers de tour est effectué autour de l'origine des coordonnées pendant le temps 2π). La dérivée de la vitesse angulaire ω par rapport au paramètre est différente de zéro dans le cas général. Donc pour paramètre nous pouvons prendre cette vitesse ou, ce qui est encore mieux, son écart par rapport à $1/3$ que nous désignerons par ε et appellerons *désaccord de fréquence*. La valeur de résonance du paramètre correspond à $\varepsilon = 0$. Ce qui nous intéresse c'est le comportement du système pour des ε petits.

Si l'on néglige les termes non linéaires dans les équations de Hamilton et le désaccord de fréquence ε, alors toutes les trajectoires

*) Le procédé mentionné ici est utile aussi bien pour l'étude des systèmes hamiltoniens que dans la théorie générale des équations différentielles ordinaires. Voir par exemple, V. Arnold, *Cours sur les bifurcations et les familles verselles* (en russe). « Ouspekhi matematitcheskikh naouk », 27, n° 5 (1972), 120-184.

de notre système se fermeront après avoir effectué trois tours complets (i. e. auront une période égale à 6π). On se propose maintenant d'étudier l'influence des termes non linéaires et du désaccord de fréquence sur le comportement des trajectoires. Il est clair que les trajectoires ne se fermeront pas toutes dans le cas général. Pour voir comment elles se comportent il faut envisager la forme normale.

Dans le système de coordonnées choisi $z = p + iq$; $\bar{z} = p - iq$ la fonction de Hamilton s'écrit

$$-2iH = -i\omega z\bar{z} + \sum_{\alpha+\beta=3} \sum_{k=-\infty}^{+\infty} h_{\alpha\beta k} z^{\alpha} \bar{z}^{\beta} e^{ikt} + \ldots,$$

où les points de suspension désignent les termes de degré supérieur au cube et où $\omega = 1/3 + \varepsilon$.

En réduisant à la forme normale on peut éliminer tous les termes au cube à l'exception de ceux dont le petit diviseur

$$\omega (\alpha - \beta) + k$$

s'annule à la résonance. Il est équivalent de dire que ces termes sont ceux qui sont constants sur une trajectoire du mouvement périodique obtenu en négligeant le désaccord de fréquence et la non-linéarité. Ces termes sont appelés *termes résonnants*. Donc pour la résonance à la fréquence $\omega = 1/3$ les termes résonnants sont ceux pour lesquels

$$\alpha - \beta + 3k = 0.$$

Donc de tous les termes au cube ne sont résonnants que $z^3 e^{-it}$ et $\bar{z}^3 e^{it}$, et nous pouvons ramener la fonction de Hamilton à la forme

$$-2iH = -i\omega z\bar{z} + hz^3 e^{-it} - \bar{h}\bar{z}^3 e^{it} + \ldots$$

(que h et $\bar{h}$ soient conjugués découle de H réel).

Remarquons que pour réduire la fonction de Hamilton à cette forme normale nous avons effectué une transformation canonique différentiable 2π-périodique en le temps, continûment dépendant du paramètre même à la résonance. Cette transformation est l'identité aux infiniment petits près du second ordre en l'écart par rapport à la trajectoire fermée (la fonction génératrice de cette transformation ne diffère de celle de l'identité que par les termes au cube).

Continuons d'étudier le comportement des solutions des équations de Hamilton. Eliminons tout d'abord dans la fonction de Hamilton tous les termes de degré supérieur à trois et envisageons les solutions du système tronqué obtenu. Ensuite nous verrons comment les termes éliminés influeront sur le comportement des trajectoires.

On simplifie l'étude du système tronqué en introduisant sur le plan de la variable complexe z un système de coordonnées en rotation uniforme à la vitesse angulaire $1/3$, i. e. on effectue le changement $z = \zeta e^{it/3}$. Pour la variable ζ on obtient alors un système autonome de hamiltonien

$$-2iH_0 = -i\varepsilon\zeta\bar{\zeta} + h\zeta^3 - \bar{h}\bar{\zeta}^3, \text{ où } \varepsilon = \omega - 1/3.$$

C'est une véritable aubaine que le système tronqué soit autonome dans le système de coordonnées en rotation. Non seulement le système complet d'équations de Hamilton (compte tenu des termes du hamiltonien de degré supérieur à trois) n'est pas autonome dans le système de coordonnées en rotation, mais il n'est même pas 2π-périodique (il est 6π-périodique) en le temps. Le système autonome de hamiltonien H_0 est en fait le résultat de la moyennisation du système initial par rapport aux trajectoires fermées du système linéaire avec $\varepsilon = 0$ (aux termes près de degré supérieur à trois).

Le coefficient h peut être rendu réel (moyennant une rotation du système de coordonnées). Donc rapportée aux coordonnées réelles (x, y), la fonction de Hamilton s'écrit

$$H_0 = \frac{\varepsilon}{2}(x^2 + y^2) + a(x^3 - 3xy^2).$$

Le coefficient a dépend du désaccord de fréquence ε comme d'un paramètre. Pour $\varepsilon = 0$ ce coefficient est en général non nul. On

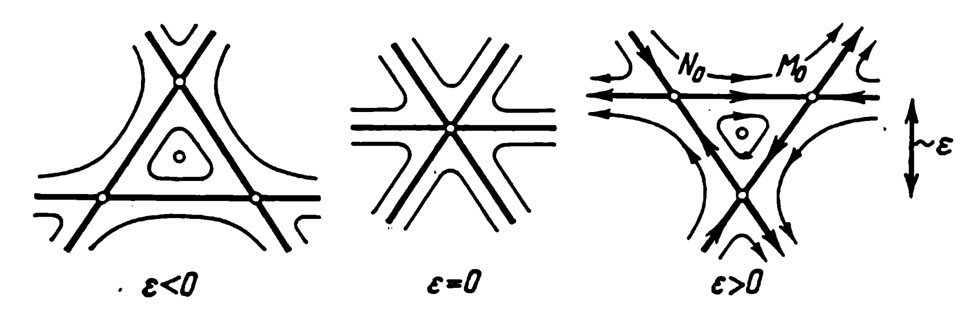


Fig. 239. Passage par la résonance 3 : 1.

peut donc le rendre égal à 1 par un changement de coordonnées dépendant différentiablement du paramètre. Ainsi il nous faut étudier la modification des courbes du système de hamiltonien

$$H_0 = \frac{\varepsilon}{2}(x^2 + y^2) + (x^3 - 3xy^2)$$

en fonction du paramètre ε dans le plan (x, y).

On voit aisément que les courbes se modifieront comme suit (fig. 239). Pour $\varepsilon = 0$ la ligne de niveau nul de la fonction H_0 est composée de trois droites qui se coupent en 0 sous des angles de 60°. Lorsque ε varie il existe toujours une ligne de niveau composée de trois droites qui se déplacent en mouvement de translation en formant un triangle équilatéral ayant l'origine des coordonnées pour centre. Les sommets de ce triangle sont des points critiques du type selle du hamiltonien. Au passage de ε par zéro (i. e. au passage par la résonance) le point critique confondu avec l'origine des coordonnées de minimum devient maximum.

Donc pour le système de fonction de Hamilton H_0 l'origine des coordonnées est une position d'équilibre stable pour toutes les valeurs du paramètre, à l'exception de la valeur de résonance, pour laquelle

l'équilibre est instable. Pour les valeurs du paramètre voisines de la résonance, le triangle qui au voisinage de l'origine des coordonnées est rempli de courbes de phase fermées, est petit (de l'ordre de ε) de sorte que le « rayon de stabilité » de l'origine des coordonnées tend vers zéro lorsque $\varepsilon \to 0$: il suffit d'une petite perturbation de la condition initiale (de l'ordre de ε) pour que le point représentatif quitte le triangle et s'éloigne de la position d'équilibre.

Formulons ces résultats pour le problème initial de la trajectoire périodique (ces résultats n'ont certes pas été démontrés puisque nous avons négligé les termes supérieurs au cube, mais ils peuvent être justifiés).

1. Au passage par la résonance $3:1$ considérée la trajectoire périodique devient en général instable.

2. Lorsque le paramètre prend des valeurs voisines de la résonance, il existe une trajectoire périodique instable au voisinage de la trajectoire périodique considérée sur la même variété de niveau d'énergie. Cette trajectoire se ferme après avoir longé trois fois et s'être enroulée une fois sur la trajectoire initiale. Ces deux trajectoires se confondent à la valeur de résonance du paramètre.

3. La distance de la trajectoire périodique instable à la trajectoire stable décroît à l'approche de la résonance comme la première puissance du désaccord de fréquence (i. e. comme la première puissance de l'écart du paramètre à la valeur de résonance).

4. Par la trajectoire instable, sur la même variété de niveau d'énergie de dimension trois, passent deux surfaces invariantes de dimension deux qui sont les adhérences des trajectoires qui s'approchent de la trajectoire périodique instable sur l'une ou l'autre de ces surfaces selon que $t \to +\infty$ ou $t \to -\infty$.

5. Les séparatrices sont disposées de telle sorte que leur intersection avec la surface transversale à la trajectoire initiale donne une figure qui rappelle les trois côtés d'un triangle équilatéral et leurs prolongements. Les sommets de ce triangle sont les points d'intersection de la trajectoire périodique instable avec la surface transversale.

6. Si les conditions initiales sont contenues dans le triangle formé par les séparatrices, le point représentatif reste assez longtemps (un temps de l'ordre d'au moins $1/\varepsilon$) au voisinage de la trajectoire périodique initiale (à une distance de l'ordre de ε). Dans le cas contraire le point s'éloigne assez rapidement à une distance grande par rapport à ε.

E. Bifurcation des séparatrices. En réalité les séparatrices mentionnées dans les propositions 4, 5 et 6 sont très compliquées (en raison de l'influence des termes de degré supérieur au cube que nous avons rejetés). Pour bien se fixer les idées il faut considérer une section σ coupant transversalement la trajectoire fermée initiale en l'un

quelconque de ses points (et contenu entièrement dans une variété de niveau d'énergie) *).

Les trajectoires issues de cette section la recoupent de nouveau au bout d'un intervalle de temps proche de la période de révolution sur la trajectoire fermée initiale. Il s'introduit donc une application d'un voisinage du point d'intersection de la trajectoire fermée avec la section considérée sur cette section. Cette application possède un point fixe (le point d'intersection de la section avec la trajectoire fermée) et est voisine d'une rotation de 120° autour de ce point que nous prendrons pour origine des coordonnées sur notre section.

Considérons maintenant le cube de l'application précédente. C'est une application d'un voisinage du zéro du plan sur la surface, laissant fixe l'origine des coordonnées. A présent cette application est voisine d'une rotation de 360°, i.e. de l'identité: elle est réalisée par les trajectoires de notre système en un intervalle de temps proche de trois périodes de la trajectoire fermée considérée.

Les calculs effectués plus haut fournissent des renseignements non triviaux sur la structure de cette « application en trois périodes ». En effet, en négligeant les termes du hamiltonien de puissance supérieure au cube, nous modifions les termes de l'application de puissance supérieure au carré. Donc l'application en trois périodes qui correspond au hamiltonien tronqué approche (avec une erreur cubique) la vraie application en trois périodes.

Or nous connaissons les propriétés de l'application en trois périodes qui correspond au hamiltonien tronqué puisque c'est une application du flot du système de hamiltonien $H_0(x, y)$ pendant 6π (la démonstration se base sur le fait que le système de coordonnées en rotation revient à sa position initiale au bout d'un temps égal à 6π). Voyons maintenant quelles sont les propriétés qui se conservent par une perturbation du troisième ordre de petitesse par rapport à la distance au point fixe.

Désignons par A_0 l'application en trois périodes correspondant au système tronqué et par A la vraie application.

1. L'application A_0 fait partie du flot: c'est une transformation pendant le temps 6π du flot de hamiltonien H_0.

Rien ne nous permet de conclure que l'application A fait partie du flot.

2. L'application A_0 se conserve par une rotation de 120°: il existe un difféomorphisme non trivial g commutant avec A_0 et tel que $g^3 = E$.

Nous n'avons aucune raison de penser que l'application A commute avec un difféomorphisme g non trivial quelconque tel que $g^3 = E$.

---

*) Ici il est plus commode de raisonner sur les applications en une période et plus facile d'effectuer les calculs sur les flots.

3. L'application A_0 possède trois points fixes instables situés à une distance de l'ordre de ε de l'origine des coordonnées et proches des sommets du triangle équilatéral. Pour ε suffisamment petit (ε représente l'écart par rapport à la résonance) l'application A possède aussi trois points fixes instables au voisinage des sommets du triangle équilatéral. Ceci découle du théorème des fonctions implicites.

4. Les séparatrices des points fixes de l'application A_0 forment (pour les valeurs du paramètre voisines de la résonance) une figure

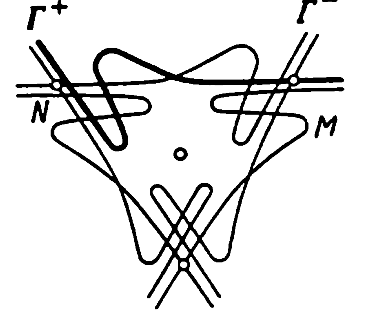


Fig. 240. Bifurcation des séparatrices.

qui rappelle les côtés d'un triangle équilatéral et leurs prolongements. Si l'on fait opérer l'application A_0 plusieurs fois en partant d'un point d'un côté, que nous désignerons par M_0N_0, du triangle (voir fig. 239), on obtient une suite de points sur M_0N_0 qui tend vers l'un des sommets, disons M_0. Si on fait opérer l'application A_0^{-1} on obtient une suite qui tend vers N_0.

Chacun des trois points fixes instables de l'application A possède des séparatrices voisines des côtés du triangle (fig. 240). Plus exactement, les points du plan qui tendent vers le point fixe M par les applications A^n, $n \to +\infty$ forment une courbe différentiable Γ^+, invariante par A, qui passe par le point M et est proche, au voisinage du point M, du côté M_0N_0 du triangle formé par les séparatrices de l'application A_0. Les points qui tendent vers N par les applications A^n où $n \to -\infty$ forment une autre courbe Γ^- qui est invariante, différentiable, passe par le point N et au voisinage de N_0 est proche de M_0N_0.

Cependant *les courbes* Γ^+ *et* Γ^- *qui sont toutes deux proches de la droite* M_0N_0 *ne sont pas nécessairement confondues.* C'est précisément en cela que consiste la bifurcation des séparatrices qui établit une distinction fondamentale entre le comportement des trajectoires du système tronqué et celles du système complet.

La grandeur de la bifurcation des séparatrices pour ε petit est exponentiellement petite, c'est pourquoi elle passe souvent inaperçue dans les calculs suivant tel ou tel schéma de la théorie des perturbations. Cependant elle est importante dans son principe même. Exemple: sa seule existence entraîne la *divergence* des séries d'une foule de variantes de la théorie des perturbations (puisque ces séries convergeraient s'il n'y avait pas cette bifurcation).

26—01408

D'une façon générale, la divergence des séries de la théorie des perturbations (pour une bonne approximation assurée par quelques premiers termes) tient au fait qu'on cherche un objet inexistant. Si nous essayons de simuler le phénomène étudié par un schéma qui en réalité n'en reproduit pas les traits essentiels il n'est pas étonnant que nos séries divergent.

Les séries de Birkhoff (que l'on obtient si dans la réduction on ne se limite pas à quelques premiers termes du développement taylorien du hamiltonien, mais l'on va jusqu'à l'infini) sont un exemple de schéma de la théorie des perturbations, un schéma qui est formellement valable mais en fait est divergent. Si ces séries étaient convergentes, le système oscillatoire général à un degré de liberté et à coefficents périodiques se ramènerait au voisinage d'une position d'équilibre à une forme autonome normale et il n'y aurait pas de bifurcation des séparatrices (pourtant elle existe).

S'agissant de la trajectoire fermée initiale, nous constatons qu'aux trois points fixes instables de l'application A correspond une trajectoire fermée instable voisine de la trajectoire initiale longée trois fois. Il existe une famille de trajectoires tendant vers cette trajectoire instable pour $t \to +\infty$ et une autre famille pour $t \to -\infty$. Les points des trajectoires de chaque famille forment une surface différentiable qui contient notre trajectoire instable.

Ces deux surfaces sont justement les séparatrices mentionnées dans les propositions 4, 5, 6 de la page 399. Leur intersection avec

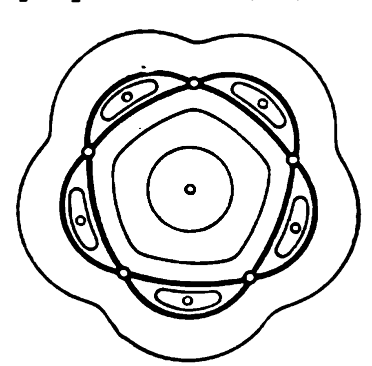


Fig. 241. Moyenne du hamiltonien des oscillations de phase du voisinage de la résonance 5 : 1.

notre surface transversale donne les courbes invariantes Γ^* et Γ^- de l'application A. En se coupant, ces deux courbes forment un réseau embrouillé dont Poincaré qui le premier a découvert le phénomène de bifurcation des séparatrices écrivait : « Que l'on cherche à se représenter la figure formée par ces deux courbes et leurs intersections en nombre infini dont chacune correspond à une solution doublement asymptotique, ces intersections forment une sorte de treillis, de tissu, de réseau à mailles infiniment serrées ; chacune des deux courbes ne doit jamais se recouper elle-même, mais elle doit se replier sur elle-même d'une manière très complexe pour venir recouper une infinité de fois toutes les mailles du réseau.

On sera frappé de la complexité de cette figure, que je ne cherche même pas à tracer. Rien n'est plus propre à nous donner une idée de la complication du problème des trois corps et en général de tous les problèmes de la dynamique où il n'y a pas d'intégrale uniforme

et où les séries de Bohlin sont divergentes. » (H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, v. 2, chap. XXXIII, page 389).

A remarquer que bien des choses ne sont pas claires dans le tableau offert par l'intersection des séparatrices.

E. Résonances d'ordre supérieur. Les résonances d'ordre supérieur peuvent être également étudiées à l'aide de la forme normale. Observons que les résonances d'ordre supérieur à quatre ne provoquent pas d'instabilité puisque la forme normale contient des termes à la puissance quatre qui assurent à la fonction H_0 un minimum ou un maximum même à la résonance.

Dans le cas d'une résonance d'ordre $n > 4$ la modification générique des courbes de phase du système de hamiltonien H_0 (fig. 241) est donnée par la formule

$$H_0 = \varepsilon\tau + \tau^2\alpha(\tau) + a\tau^{n/2}\sin n\varphi,$$
$$2\tau = p^2 + q^2, \quad \alpha(0) = \pm 1$$

Pour un petit écart (de l'ordre de ε) de la fréquence par rapport à la résonance et à une petite distance (de l'ordre de $\sqrt{|\varepsilon|}$) de la position d'équilibre (qui est placée en l'origine des coordonnées) la fonction H_0 possède $2n$ points critiques au voisinage des sommets d'un $2n$-polygone régulier ayant l'origine des coordonnées pour centre. La moitié de ces points critiques sont du type selle, l'autre des maximums si l'origine des coordonnées est un minimum et des minimums si l'origine des coordonnées est un maximum. Les points selles et les points stables alternent. Tous les n points selles sont situés sur un niveau de la fonction H_0 et les séparatrices qui joignent des selles successives forment n « îlots » dont chacun est une adhérence de courbes de phase fermées entourant le point stable. Les îlots ont une largeur de l'ordre de $\varepsilon^{(n/4)-1}$. Les courbes de phase fermées contenues dans chaque îlot sont appelées *oscillations de phase* (puisqu'en principe c'est la phase des oscillations autour de l'origine des coordonnées qui varie). La période des oscillations de phase croît comme $\varepsilon^{-n/4}$ lorsque le désaccord de fréquence ε diminue.

L'anneau étroit formé par les îlots contient des courbes de phase fermées qui encerclent l'origine des coordonnées; à l'extérieur de cet anneau les courbes de phase sont aussi fermées mais elles ne sont pas parcourues dans le même sens qu'à l'intérieur. On remarquera que le rayon de l'anneau est de l'ordre de $\sqrt{|\varepsilon|}$ indépendamment de l'ordre de la résonance, pourvu que celui-ci soit supérieur à 4. Ceci étant, l'anneau n'existe que pour l'un des deux signes de ε.

Lorsqu'on passe du système tronqué de hamiltonien H_0 au système complet les séparatrices bifurquent comme nous l'avons décrit pour la résonance d'ordre 3. L'écart des séparatrices est exponentielle-

ment petit (de l'ordre de $e^{-1/\varepsilon^{n/4}}$), cependant cette bifurcation a une importance de principe dans l'étude de la stabilité surtout dans le cas multidimensionnel.

S'agissant de notre trajectoire fermée initiale, nous obtenons le tableau suivant. Lorsqu'on se rapproche de la résonance suivant l'axe ε et d'un côté bien défini *), deux trajectoires, l'une stable et l'autre instable, se séparent de la trajectoire périodique. Ces deux nouvelles trajectoires se ferment après avoir effectué n tours le long de la trajectoire initiale dont elles sont séparées par une distance de l'ordre de $\sqrt{|\varepsilon|}$. Au voisinage de la trajectoire stable il existe une zone d'oscillations de phase lentes de période de l'ordre de $\varepsilon^{-n/4}$ et d'amplitude de l'ordre de π/n dans la direction azimutale et de l'ordre de $\varepsilon^{(n/4)-1}$ dans la radiale. La trajectoire périodique initiale ne perd pas sa stabilité au passage par la résonance, du moins dans l'approximation considérée.

Le cas de la résonance du quatrième ordre fait un peu exception. En effet, dans ce cas la forme normale contient des termes résonnants et des termes non résonnants à la puissance quatre. L'allure des courbes de phase du système tronqué dépend des termes prépondérants. Si ce sont les termes résonnants, la modification est la même que pour la résonance d'ordre trois sauf que le triangle sera remplacé par un carré. Si ce sont les termes non résonnants la modification est la même que pour $n > 4$.

Remarquons en conclusion que l'approximation fournie par la forme normale est d'autant meilleure que nous sommes près de la résonance ($\varepsilon \ll 1$) et que l'écart du point initial par rapport à la trajectoire périodique est petit. Plus exactement, lorsqu'on se rapproche d'un quotient exact de la période de la trajectoire fermée par la période d'oscillations des trajectoires voisines et que la condition initiale soit proche de la trajectoire fermée, alors augmente le temps sur lequel notre approximation décrit correctement l'allure des courbes de phase.

Ces raisonnements ne nous permettent de tirer aucune conclusion sur le comportement des courbes de phase non fermées sur un intervalle de temps infini (par exemple, sur la stabilité au sens de Liapounov de la trajectoire périodique initiale) puisque les termes négligés dans la réduction à la forme normale peuvent totalement modifier le caractère du mouvement pendant un temps infini. En réalité dans les conditions envisagées la trajectoire périodique initiale est stable au sens de Liapounov, mais la démonstration fait appel à un nouvel outil plus puissant que la forme normale de Birkhoff (voir Appendice 8).

---

*) Contrairement à la résonance d'ordre 3 pour laquelle la trajectoire périodique instable bifurquée existe des deux côtés de la résonance.

APPENDICE 8

THÉORIE DES PERTURBATIONS DES MOUVEMENTS QUASI PÉRIODIQUES ET THÉORÈME DE KOLMOGOROV

Le nombre de problèmes exactement « intégrables » n'est pas énorme (problèmes de dimension un, mouvement d'un point dans un champ central, mouvements eulérien et lagrangien du solide, problème de deux centres fixes, mouvement sur les géodésiques d'un ellipsoïde). Cependant ces « cas intégrables » nous permettent de recueillir une foule d'information sur le mouvement de systèmes plus importants en prenant le problème intégrable pour première approximation.

Exemple : le mouvement des planètes autour du Soleil. La masse des planètes étant environ mille fois moindre que celle du Soleil, dans une première approximation on pourra négliger l'interaction des planètes et ne prendre en considération que leur attraction par le Soleil. On obtient donc un problème exactement intégrable ; chaque planète indépendamment des autres décrira son ellipse keplérienne et le mouvement du système sera globalement quasi périodique. Si maintenant l'on tient compte de l'interaction des planètes, alors le mouvement keplérien de chacune d'elles variera.

La théorie des perturbations de la mécanique céleste se devra de tenir compte de cette interaction.

Il est clair que les calculs effectués sur une période de l'ordre de 1 000 ans ne doivent pas en principe poser de difficultés. Cependant si nous voulons étudier des intervalles de temps plus grands et en particulier si nous nous intéressons au comportement des solutions exactes des équations de mouvement sur un intervalle de temps infini alors ces difficultés ne manqueront pas de surgir.

En effet, l'accumulation des perturbations sur un intervalle de temps grand devant 1 000 ans peut complètement modifier le caractère du mouvement : les planètes peuvent s'écraser sur le Soleil, s'en éloigner et se heurter.

Remarquons que le comportement des solutions des équations de mouvement sur des intervalles de temps infinis n'a qu'un rapport indirect avec le mouvement des planètes. En effet, sur des intervalles de temps de l'ordre du milliard on ressent fortement les effets non conservatifs petits qui ont été négligés dans les équations de Newton. Donc les effets de l'interaction gravitationnelle ne sont

essentiels que s'ils modifient sérieusement le caractère du mouvement sur un intervalle de temps fini, petit devant le temps que mettent les effets non conservatifs à se manifester.

Ces mouvements se prêtent bien aux calculs sur ordinateurs : on peut connaître leur passé et leur futur sur des milliers d'années.

Cependant il importe de remarquer que même les ordinateurs les plus perfectionnés sont inaptes à prédire l'influence des perturbations si le point représentatif tombe dans la zone d'instabilité exponentielle.

Les méthodes asymptotiques et qualitatives possèdent une importance plus grande encore dans l'étude du mouvement de particules chargées dans les champs magnétiques puisque la particule est plus rapide que la calculatrice et effectue un nombre si élevé de tours que la détermination de sa trajectoire est impossible même en l'absence d'instabilité exponentielle.

On a élaboré de nombreuses méthodes qui tiennent compte des perturbations en mécanique céleste. (Elles sont traitées en détail dans l'ouvrage de H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*. Œuvres choisies. Dover publications inc., New York, 1957). Le trait commun de ces méthodes est qu'elles conduisent à des séries convergentes et ne fournissent par conséquent aucun renseignement sur le comportement du mouvement global sur des intervalles de temps infinis.

La divergence des séries de la théorie des perturbations est due à la présence des « petits diviseurs » : des combinaisons linéaires entières de fréquences des mouvements non perturbés par lesquelles il faut diviser pour calculer l'influence des perturbations. Dans le cas de la résonance exacte (i. e. lorsque les fréquences sont commensurables) ces diviseurs s'annulent et le terme correspondant de la série devient infiniment grand. Au voisinage de la résonance ce terme devient très grand.

Ainsi dans leur mouvement autour du Soleil Jupiter et Saturne parcourent respectivement 299 et 120,5 secondes d'arcs par jour. Donc le diviseur $2\omega_J - 5\omega_S$ est très petit devant chacune de ces fréquences. Ceci conduit à une perturbation mutuelle des planètes étalée sur une longue période (de l'ordre de 8 siècles) : l'étude de cet effet par Laplace fut l'un des premiers succès de la théorie des perturbations

Remarquons que la difficulté posée par les petits diviseurs est essentielle. En effet, les nombres rationnels forment un ensemble partout dense. Donc dans l'espace des phases du problème non perturbé, les conditions initiales pour lesquelles il y a résonance et pour lesquelles les petits diviseurs s'annulent forment un ensemble partout dense. Par conséquent les fonctions auxquelles conduisent les séries de la théorie des perturbations possèdent un ensemble de points singuliers partout dense.

Cette difficulté n'est pas inhérente aux seuls problèmes de mécanique céleste, elle caractérise également tous problèmes proches des intégrables (par exemple, le mouvement de la toupie pesante asymétrique en rotation rapide). Poincaré a même qualifié de *problème particulier de la dynamique* l'étude des perturbations des mou-

vements quasi périodiques d'un système défini par le hamiltonien

$$H = H_0(I) + \varepsilon H_1(I, \varphi), \quad \varepsilon \ll 1,$$

dans les variables action-angle I, φ.

H_0 représente ici le hamiltonien du problème non perturbé, εH_1 la perturbation qui est une fonction 2π-périodique des variables angulaires $\varphi_1, \ldots, \varphi_n$. Dans le problème non perturbé ($\varepsilon = 0$) les angles φ varient uniformément avec des fréquences constantes

$$\omega_k = \partial H_0 / \partial I_k;$$

quant aux variables action, elles sont toutes intégrales premières.

Soit à étudier les courbes de phase des équations de Hamilton

$$\dot{I} = -\partial H / \partial \varphi, \quad \dot{\varphi} = \partial H / \partial I$$

dans l'espace des phases défini par le produit cartésien d'un domaine d'un espace de dimension n muni des coordonnées I par un tore de dimension n muni des coordonnées angulaires φ.

Un important progrès a été fait dans l'étude des courbes de phase perturbées de ce problème grâce au travail de A. Kolmogorov, *Sur la conservation des mouvements quasi périodiques pour une faible variation de la fonction de Hamilton* (en russe), DAN, **98**, n° 4 (1954), 527-530. Dans le présent appendice on expose les principaux résultats obtenus depuis dans ce domaine. Les démonstrations figurent dans les travaux suivants:

1. V. Arnold — Petits diviseurs I. Sur les applications du cercle sur lui-même (en russe). « Izvestia Academii Naouk de l'U.R.S.S. », **25**, n° 1 (1961), 21-86.

2. V. Arnold — Petits diviseurs II. Démonstration du théorème de A. Kolmogorov sur la conservation des mouvements quasi périodiques pour une faible variation de la fonction de Hamilton (en russe). « Ouspekhi matematitcheskikh naouk », **18**, n° 5 (1963), 13-40.

3. V. Arnold — Petits diviseurs III. Petits diviseurs et problèmes de stabilité dans la mécanique classique et la mécanique céleste (en russe). « Ouspekhi matematitcheskikh naouk », **18**, n° 6 (1963), 81-192.

4. V. Arnold, A. Avez — Problèmes ergodiques de la mécanique classique. Paris, G. V., 1967.

5. J. K. Moser — On invariant curves of area-preserving mapping of an annulus. Nachr. Acad. Wiss. Göttingen, Math. Phys. K1 **11a**, n° 1 (1962), 1-20.

6. J. K. Moser — A rapidly convergent methode and nonlinear differential equations. Annali della Scuola Norm. Super. de Pisa, ser. III, **20**, n° 2 (1966), 265-315, n° 3 (1966), 499—535.

7. J. K. Moser — Convergent series expansions for quasiperiodic motions. Math. Ann. **169** (1967), 136-176.

8. C. L. S i e g e l, J. K. M o s e r — Lectures on Celestial Mechanics. Springer, 1971.
9. S. S t e r n b e r g — Celestial Mechanics I, II. N.Y., Benjamin, 1969.

Avant de formuler ces résultats nous allons brièvement discuter le comportement des courbes de phase du problème non perturbé traité dans le chapitre 10.

A. Mouvement non perturbé. Un système de hamiltonien $H_0(I)$ possède n intégrales premières en involution (n variables action). Tout ensemble de niveau de ces intégrales est un tore de dimension n dans l'espace des phases de dimension $2n$. Ce tore est invariant par le flot du système non perturbé : toute courbe de phase issue d'un point d'un tel tore reste sur ce tore.

Le mouvement du point représentatif sur un tore invariant $I = \text{const}$ est quasi périodique. Les fréquences de ce mouvement sont les dérivées du hamiltonien non perturbé par rapport aux variables action :

$$\dot{\varphi}_k = \omega_k(I), \text{ où } \omega_k = \partial H_0/\partial I_k.$$

Donc la courbe de phase est partout dense sur un tore dont la dimension est égale au nombre de fréquences ω_k rationnellement indépendantes.

Remarquons que les fréquences dépendent du tore étudié, i.e. des valeurs des intégrales premières que nous avons fixées. D'une façon générale, le système des n fonctions $\omega(I)$ est fonctionnellement indépendant ; dans ce cas nous pouvons tout simplement numéroter les tores avec les fréquences en prenant les variables ω pour coordonnées au voisinage du point considéré dans l'espace des variables action I.

Lorsque les fréquences sont fonctionnellement dépendantes on dit qu'on a affaire au *cas non dégénéré*. Donc les conditions de non-dégénérescence s'écrivent

$$\det\left|\frac{\partial\omega}{\partial I}\right| = \det\left|\frac{\partial^2 H_0}{\partial I^2}\right| \neq 0.$$

Ainsi, dans le cas non dégénéré les mouvements sont quasi périodiques à nombre de fréquences différent sur les divers tores invariants de l'espace des phases du problème non perturbé. En particulier, les tores invariants sur lesquels le nombre de fréquences est maximal (i. e. égal à n) forment un ensemble partout dense dans l'espace des phases. Ces tores sont appelés *non résonnants*.

On montre que les tores non résonnants forment dans l'espace des phases un ensemble de mesure complète, de sorte que la mesure de Lebesgue de l'union de tous les tores invariants résonnants du système non dégénéré non perturbé est nulle. Pourtant les tores

résonnants invariants existent et alternent avec les tores non résonnants de telle sorte qu'ils forment également un ensemble partout dense. Bien plus, est partout dense l'ensemble des tores résonnants à nombre quelconque de 1 à $n - 1$ fréquences indépendantes. En particulier, les tores invariants sur lesquels toutes les courbes de phase sont fermées (le nombre de fréquences indépendantes est 1) forment un ensemble partout dense.

Remarquons tout de même que la probabilité de tomber sur un tore résonnant si le point initial est aléatoirement choisi dans l'espace des phases du système non perturbé est nulle (tout comme la proba-

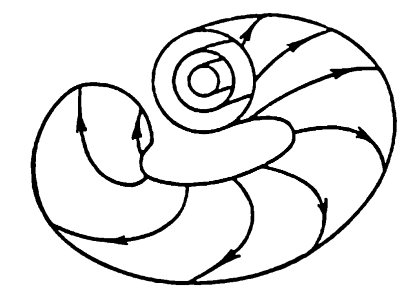


Fig. 242. Tores invariants dans une variété de dimension trois de niveau d'énergie.

bilité de choisir un nombre rationnel parmi les réels). Donc en négligeant les ensembles de mesure nulle on peut dire que presque tous les tores invariants du système non perturbé non dégénéré sont non résonnants et possèdent une collection complète de n fréquences rationnellement indépendantes.

La trajectoire d'un mouvement quasi périodique est partout dense sur un tore non résonnant. Donc pour presque toutes les conditions initiales la courbe de phase d'un système non perturbé non dégénéré est partout dense sur un tore dont la dimension est égale au nombre de degrés de liberté (i. e. la moitié de la dimension de l'espace des phases).

Illustrons ce que nous venons de dire sur le cas de deux degrés de liberté ($n = 2$). L'espace des phases est de dimension quatre. Donc les ensembles de niveau d'énergie sont de dimension trois. Fixons l'un d'eux. C'est une variété de dimension trois, feuilletée en tores invariants de dimension deux que l'on peut se représenter dans l'espace ordinaire comme une famille de tores concentriques plongés les uns dans les autres (fig. 242).

Les courbes de phase sont les hélices de ces tores ; en outre les deux fréquences de révolution changent d'un tore à l'autre. Dans le cas général non seulement les deux fréquences mais leur rapport varieront d'un tore à l'autre. Si la dérivée du rapport des fréquences par rapport à la variable action qui numérote les tores sur l'ensemble de niveau donné de H_0 est non nulle, nous dirons que notre système est *isoénergétiquement non dégénéré*. La condition de non-dégéneres-

cence isoénergétique s'écrit (comme le montre un calcul simple)

$$\det \begin{vmatrix} \frac{\partial^2 H_0}{\partial I^2} & \frac{\partial H_0}{\partial I} \\ \frac{\partial H_0}{\partial I} & 0 \end{vmatrix} \neq 0.$$

Les conditions de non-dégénérescence et de non-dégénérescence isoénergétique sont indépendantes l'une de l'autre i.e. l'une n'entraîne pas l'autre. Dans le cas multidimensionnel ($n > 2$) la non-dégénérescence isoénergétique signifie la non-dégénérescence de l'application suivante de la variété de dimension $(n - 1)$ de niveau de la fonction $H_0(I)$ des n variables action dans l'espace projectif de dimension $(n - 1)$:

$$I \longmapsto (\omega_1(I) : \omega_2(I) : \ldots : \omega_n(I)).$$

Soit donc à étudier un système non dégénéré isoénergétique à deux degrés de liberté. Il est aisé de construire une petite section coupant transversalement les 2-tores de notre famille (en s'inspirant d'une famille de cercles concentriques dans l'espace ordinaire).

Une courbe de phase issue de cette section y reviendra après avoir accompli un tour de tore. Nous obtenons ainsi un point sur le cercle suivant lequel le tore coupe cette section. Il s'introduit donc une application de cette section sur elle-même.

Cette application laisse invariants les cercles méridiens concentriques suivant lesquels les tores invariants coupent cette section. Ceci étant, tout cercle tourne d'un certain angle, plus exactement d'une fraction de tour égale au quotient de la fréquence de rotation sur un méridien par la fréquence de rotation sur l'équateur. Donc si un système n'est pas isoénergétiquement dégénéré, l'angle de rotation des cercles invariants sur le plan d'intersection changera d'un cercle à l'autre.

Donc sur certains cercles cet angle sera commensurable avec un tour complet et sur d'autres non. Les uns et les autres formeront un ensemble partout dense, mais sur presque tous les cercles (au sens de la mesure de Lebesgue) l'angle de rotation sera incommensurable avec un tour complet.

La commensurabilité et l'incommensurabilité influent de la manière suivante sur le comportement des points d'un cercle par l'application de la sectoin sur elle-même. Si l'angle de rotation est commensurable avec un tour complet, alors après quelques itérations de l'application le point revient à sa position initiale (le nombre d'itérations est d'autant plus grand que le dénominateur de la fraction exprimant cet angle de rotation est grand). Si l'angle de rotation est incommensurable avec un tour complet, alors les images successives du point par l'application sont partout denses sur le cercle méridien.

Remarquons encore que la commensurabilité correspond aux tores résonnants, et l'incommensurabilité aux non résonnants. Re-

marquons également que l'existence des tores résonnants entraîne le fait suivant. Considérons une puissance de l'application de notre section sur elle-même. Supposons que l'exposant de cette puissance est le dénominateur du rapport de fréquences sur l'un des tores résonnants. Alors l'application portée à la puissance indiquée possède un cercle entièrement composé de points fixes (plus exactement un méridien du tore résonnant considéré).

Un tel comportement des points fixes est contre nature pour les applications d'une forme tant soit peu générique, même pour les canoniques (en général les points fixes sont isolés). Ici le cercle entier de points fixes résulte du fait que nous considérons un système intégrable non perturbé. Pour une perturbation générique aussi petite que l'on veut la propriété de l'application (de posséder tout un cercle de points fixes) doit disparaître. Le cercle de points fixes doit se désagréger et de tous ses points il n'en restera qu'un nombre fini.

En d'autres termes, pour une faible perturbation du système intégrable il faut s'attendre à voir varier l'allure des courbes de phase ne serait-ce que parce que les tores invariants qui sont les adhérences des courbes de phase fermées doivent se décomposer et il ne restera qu'un nombre fini de courbes fermées voisines des courbes non perturbées, les autres ayant un comportement plus compliqué. Nous avons déjà vu un cas analogue dans l'Appendice 7 lorsque nous avons étudié les oscillations de phase au voisinage de la résonance.

Voyons maintenant ce qui se passe avec les tores invariants non résonnants pour une faible perturbation du hamiltonien. L'application formelle du principe de la moyenne (i.e. la première approximation de la théorie classique des perturbations, cf. § 52) nous amène à conclure que le *tore non résonnant ne subit aucune évolution.*

Remarquons qu'ici la hamiltonicité des perturbations est essentielle puisque pour une perturbation non conservative les variables d'action peuvent évidemment évoluer. Si l'on transpose ceci à la mécanique céleste, l'évolution des variables d'action se traduit par une variation séculaire des demi grands axes des ellipses de Kepler, i.e. par la chute des planètes sur le Soleil ou leur collision, ou leur fuite à une grande distance pendant un temps inversement proportionnel à la grandeur de la perturbation. Si les perturbations conservatives conduisaient à une évolution en première approximation, cela ne manquerait pas de se répercuter sur le destin des planètes dans un espace de temps de l'ordre de 1 000 ans. Heureusement l'ordre de grandeur des perturbations non conservatives est de beaucoup plus petit.

Le théorème de Kolmogorov formulé plus bas justifie la conclusion de la théorie non rigoureuse des perturbations relative à l'absence d'évolution des variables d'action.

B. Tores invariants du système perturbé.

Théorème. *Si un système hamiltonien est non dégénéré, alors la plupart des tores invariants non résonnants ne disparaissent pas par une*

perturbation hamiltonienne conservative suffisamment petite, mais se déforment légèrement de telle sorte que dans l'espace des phases du système perturbé il existe également des tores invariants, adhérences de courbes de phase qui en sont des hélices quasi périodiques à nombre de fréquences égal au nombre de degrés de liberté.

Les tores invariants indiqués sont la majorité en ce sens que la mesure du complémentaire de leur union est petite avec la perturbation.

La démonstration du théorème de Kolmogorov se base sur les deux remarques suivantes.

1. Fixons une collection non résonnante de fréquences du système non perturbé telle que ces fréquences soient non seulement indépendantes mais ne satisfassent approximativement à aucune relation de résonance de petit ordre.

Plus exactement, fixons une collection de fréquences ω telle qu'il existe des C et ν vérifiant la condition $|(\omega, k)| > C|k|^{-\nu}$ pour tous les vecteurs entiers $k \neq 0$.

On montre que si ν est suffisamment grand (disons $\nu = n + 1$), alors la mesure de l'ensemble des vecteurs ω (contenus dans le domaine borné fixé) pour lesquels la condition de non-résonance n'est pas réalisée est petite avec C.

Ensuite, au voisinage du tore invariant non résonnant du système non perturbé, qui correspond aux valeurs fixées des fréquences, nous allons chercher un tore invariant du système perturbé sur lequel s'effectue un mouvement quasi périodique exactement avec les mêmes fréquences que celles que nous avons fixées, fréquences qui, par conséquent, satisferont la condition de non-résonance.

Donc au lieu de la variation de fréquence que l'on utilise habituellement dans de nombreux schémas de la théorie des perturbations (et qui consiste à introduire des fréquences dépendant de la perturbation) il faut que les fréquences soient non résonnantes et constantes et que les conditions initiales pour la perturbation donnée soient telles que le mouvement s'accomplisse avec des fréquences fixes. On peut réaliser cette faible variation des conditions initiales avec la perturbation, car les fréquences varient avec les variables d'action en vertu de la condition de non-dégénérescence.

2. La deuxième remarque est que, lorsqu'on cherche le tore invariant, au lieu des développements en série suivant les puissances de la perturbation qui sont couramment utilisés dans de nombreux schémas de la théorie des perturbations, on peut se servir d'une méthode à convergence rapide, telle la méthode des tangentes de Newton.

La méthode de Newton de recherche des racines d'une équation algébrique avec, au départ, une erreur ε donne après n approximations une erreur de l'ordre de ε^{2^n}. Cette *superconvergence* permet de neutraliser l'effet des petits diviseurs qui apparaissent à chaque approximation, si bien qu'il est possible non seulement d'effectuer

une infinité d'approximations, mais encore de démontrer la convergence de toute la procédure.

Les conditions qui doivent être remplies sont d'un côté l'analycité et la non-dégénérescence du hamiltonien non perturbé $H_0 (I)$ et de l'autre la non-dégénérescence et la 2π-périodicité en les variables angulaires φ du hamiltonien perturbé $\varepsilon H_1 (I, \varphi)$. La présence du petit paramètre ε n'est pas essentielle : ce qui compte c'est que la perturbation soit suffisamment petite dans un voisinage complexe quelconque de rayon ρ du plan réel des variables φ (inférieure à une fonction positive $M (\rho, H_0)$).

J. Moser a montré que la condition d'analycité peut être remplacée par une différentiabilité d'ordre suffisamment grand si l'on combine la méthode de Newton avec le lissage des fonctions à chaque approximation proposée par J. Nash.

Les mouvements quasi périodiques du système perturbé à fréquences ω fixées obtenus par cette procédure s'avèrent être des fonctions différentiables (analytiques dans le cas analytique) du paramètre ε. Donc on aurait pu les chercher sous forme d'une série entière de ε sans recourir à la méthode de Newton. Les coefficients de cette série qu'on appelle série de Lindstedt peuvent effectivement être calculés ; cependant on ne peut en démontrer la convergence qu'indirectement à l'aide des approximations newtoniennes.

C. Zones d'instabilité. La présence des tores invariants dans l'espace des phases du problème perturbé signifie que pour la plupart des conditions initiales, le mouvement d'un système voisin d'un système intégrable reste quasi périodique avec un nombre maximal de fréquences.

Et l'on se demande tout naturellement ce qu'il arrive aux autres courbes de phase dont les conditions initiales sont contenues dans les failles qui se sont constituées entre les tores invariants à la place des tores invariants résonnants du problème non perturbé.

La décomposition du tore résonnant sur lequel le nombre de fréquences est d'une unité inférieur au nombre total est facile à analyser en première approximation de la théorie des perturbations. Pour cela il faut prendre la moyenne de la perturbation sur les $(n-1)$-tores invariants en lesquels se décompose le tore invariant résonnant et sur lesquels sont partout denses les courbes de phase du système non perturbé. Cette moyennisation nous donne un système conservatif à un degré de liberté qui est facile à étudier (voir l'étude des oscillations de phase au voisinage de la résonance dans l'Appendice 7).

Dans l'approximation considérée nous obtenons, au voisinage du n-tore résonnant décomposé, une alternance de $(n-1)$-tores instables et stables, en outre des oscillations de phase ont lieu autour des tores stables. Les mouvements quasi périodiques qui leur correspondent possèdent une collection complète de n fréquences, composée de

$n-1$ fréquences rapides des oscillations initiales et d'une fréquence lente (de l'ordre de $\sqrt{\varepsilon}$) des oscillations de phase.

Cependant il faut se garder de penser que la différence qui existe entre les mouvements du système non perturbé et ceux du système perturbé se réduit à l'apparition des « îlots » d'oscillations de phase. En réalité ce phénomène est bien plus compliqué qu'il n'est décrit plus haut en première approximation. Les courbes de phase du problème perturbé ont un comportement complexe dont une illustration est la bifurcation des séparatrices traitée dans l'Appendice 7.

Lorsqu'on étudie les mouvements du système perturbé à l'extérieur des tores invariants il faut faire une distinction entre le cas de deux et celui d'un nombre plus grand de degrés de liberté. Dans le cas de deux degrés de liberté, l'espace des phases est de dimension quatre et la variété de niveau d'énergie de dimension trois. C'est pourquoi les 2-tores invariants partagent l'ensemble de niveau d'énergie.

Ceci étant, une courbe de phase qui prend naissance dans la faille entre deux tores invariants du système perturbé y reste indéfiniment. Donc quelque complexe que soit le comportement de cette courbe, elle ne quittera pas cette faille et les variables d'action correspondantes resteront indéfiniment au voisinage de leurs valeurs initiales.

Si le nombre n de degrés de liberté est supérieur à deux, alors les n-tores invariants ne partagent pas la variété de dimension $(2n-1)$ de niveau d'énergie, mais ils y seront disposés comme des points dans le plan ou des lignes dans l'espace. Dans ce cas les « failles » correspondant aux diverses résonances se joignent l'une à l'autre de sorte que les tores invariants ne font plus obstacle à l'éloignement d'une courbe de phase qui prend naissance au voisinage de la résonance. Donc nous n'avons aucune raison d'espérer que les variables d'action définies le long d'une telle courbe de phase resteront indéfiniment au voisinage de leurs conditions initiales.

En d'autres termes, dans les systèmes à deux degrés de liberté (vérifiant la condition de non-dégénérescence isoénergétique, qui en général, est réalisée), pour des perturbations suffisamment petites, les variables action définies le long d'une trajectoire de phase non seulement ne possèdent pas de perturbation séculaire dans aucune approximation de la théorie des perturbations (i.e. varient peu sur un intervalle de temps de l'ordre de $(1/\varepsilon)^N$ quel que soit N, où ε est la grandeur de la perturbation), mais restent indéfiniment au voisinage de leurs valeurs initiales aussi bien pour les courbes de phase non résonnantes quasi périodiques partout denses dans des 2-tores (et qui occupent la plus grande partie de l'espace des phases) que pour les autres conditions initiales.

Dans le même temps il existe des systèmes à nombre de degrés de liberté plus grand que deux qui vérifient toutes les conditions de non-dégénérescence, dans lesquels les variables d'action peuvent

s'éloigner lentement de leurs valeurs initiales pour certaines conditions initiales bien que le mouvement soit quasi périodique pour la plupart des conditions initiales. La vitesse moyenne de cet éloignement dans les exemples connus *) est de l'ordre de $e^{-1/\sqrt{\varepsilon}}$, i.e. elle diminue plus vite que n'importe quelle puissance de la perturbation ε. Aussi n'est-il pas étonnant que cet éloignement passe inaperçu dans toute approximation de la théorie des perturbations (par vitesse moyenne on entend ici le rapport de l'accroissement des variables d'action au temps, de sorte qu'en fait il est question d'un accroissement de l'ordre de 1 pendant un intervalle de temps de l'ordre de $e^{1/\sqrt{\varepsilon}}$).

Une majoration de la vitesse moyenne avec laquelle les variables action s'éloignent des conditions initiales dans les systèmes généraux d'équations canoniques de Hamilton à n degrés de liberté, voisins de systèmes intégrales, est faite par N. Nékhorochev **).

Cette majoration, de même d'ailleurs que la minoration citée plus haut, est de la forme e^{-1/ε^d}; donc l'accroissement des variables d'action est petit tant que le temps est petit devant e^{1/ε^d} si $\varepsilon < \varepsilon_0$. Ici ε représente la grandeur de la perturbation, d un nombre compris entre 0 et 1 et défini comme ε_0 par les propriétés du hamiltonien non perturbé H_0. En outre H_0 doit remplir une condition de non-dégénérescence (qui est longue à formuler mais qui est réalisée en général; en particulier, il suffit d'exiger la convexité quadratique du hamiltonien non perturbé H_0, i.e. que la différentielle seconde de la fonction H_0 soit définie positive ou négative).

La majoration indiquée montre que les variations séculaires des variables d'action ne sont mises en évidence par aucune approximation de la théorie des perturbations, puisque leur vitesse moyenne est exponentiellement petite. Remarquons encore que ces variations séculaires n'obéissent visiblement à aucune loi, mais sont un errement plus ou moins aléatoire entre les résonances autour des tores invariants. Les problèmes soulevés ici sont traités en détail dans l'article de G. Z a s l a v s k i et B. T c h i r i k o v, *Instabilité stochastique des oscillations linéaires* (en russe), «Ouspekhi matematitcheskih naouk», v. 105, nº 1 (1971), 3-39.

D. Diverses variantes du théorème relatif aux tores invariants. Des propositions présentant des affinités avec le théorème relatif à l'invariance des tores dans un système autonome ont été démontrées pour des équations non autonomes à coefficients périodiques et pour des applications symplectiques.

---

*) Voir B. A r n o l d, *Instabilité des systèmes dynamiques à plusieurs degrés de liberté* (en russe). DAN de l'U. R. S. S., 156, nº 1 (1964), 9-12.

**) N. N é k h o r o c h e v, *Sur le comportement des systèmes hamiltoniens proches de systèmes intégrables* (en russe). « Analyse fonctionnelle et ses applications », v. 5, fasc. 4 (1971), 82-83.

On retrouve des propositions analogues dans la théorie des petites oscillations au voisinage d'une position d'équilibre d'un système autonome ou d'un système à coefficients périodiques, de même qu'au voisinage d'une courbe de phase fermée ou au voisinage d'un point fixe d'une application symplectique.

Les conditions de non-dégénérescence diffèrent suivant les cas considérés. C'est pourquoi nous les citerons à titre d'information. Nous nous limiterons aux plus simples conditions de non-dégénérescence, à celles qui sont réalisées dans les systèmes génériques. Ces conditions peuvent souvent être allégées, mais le gain ne vaut pas la complication de l'énoncé.

1. Système autonome. Fonction de Hamilton

$$H = H_0(I) + \varepsilon H_1(I,\varphi), \quad I \in G \subset \mathbf{R}^n, \quad \varphi \bmod 2\pi \in T^n.$$

La condition de non-dégénérescence

$$\det \left| \frac{\partial^2 H_0}{\partial I^2} \right| \neq 0$$

garantit la conservation *) de la plupart des tores invariants par une petite perturbation ($\varepsilon \ll 1$).

La condition de non-dégénérescence isoénergétique

$$\det \begin{vmatrix} \frac{\partial^2 H_0}{\partial I^2} & \frac{\partial H_0}{\partial I} \\ \frac{\partial H_0}{\partial I} & 0 \end{vmatrix} \neq 0$$

garantit l'existence sur chaque variété de niveau d'énergie d'un ensemble de tores invariants dont le complémentaire possède une mesure petite. Sur ces tores les fréquences dépendent en général de la valeur de la perturbation, mais leurs rapports sont constants lorsque ε varie.

Si $n = 2$, la condition de non-dégénérescence isoénergétique assure également la stabilité des variables d'action en ce sens que celles-ci restent indéfiniment au voisinage de leurs valeurs initiales pour une perturbation suffisamment petite.

2. Système périodique. Fonction de Hamilton

$$H = H_0(I) + \varepsilon H_1(I, \varphi, t),$$
$$I \in G \subset \mathbf{R}^n, \ \varphi \bmod 2\pi \in T^n;$$

la perturbation est 2π-périodique non seulement en φ mais en t également. Il est naturel d'envisager le système non perturbé dans l'espace des phases de dimension $(2n + 1)$: $\{(I, \varphi, t)\} = \mathbf{R}^n \times T^{n+1}$. Les tores invariants sont de dimension $(n + 1)$. La condition de non-

*) Il est bien entendu que les tores se déforment légèrement par cette perturbation.

dégénérescence

$$\det \left| \frac{\partial^2 H_0}{\partial I'^2} \right| \neq 0$$

garantit la conservation de la plupart des $(n + 1)$-tores invariants par une petite perturbation $(\varepsilon \ll 1)$.

Si $n = 1$, alors cette condition de non-dégénérescence assure également la stabilité de la variable d'action en ce sens qu'elle reste indéfiniment au voisinage de sa valeur initiale pour une perturbation suffisamment petite.

3. Application $(I, \varphi) \mapsto (I', \varphi')$ d'une « $2n$-couronne ». Fonction génératrice

$$S(I', \varphi) = S_0(I') + \varepsilon S_1(I', \varphi),\ I' \in G \subset \mathbf{R}^n,\ \varphi \in T^n.$$

La condition de non-dégénérescence

$$\det \left| \frac{\partial^2 S_0}{\partial I'^2} \right| \neq 0,$$

garantit la conservation de la plupart des tores invariants de l'application non perturbée $((I, \varphi) \mapsto (I, \varphi + \partial S_0/\partial I))$ par une petite perturbation $(\varepsilon \ll 1)$.

Si $n = 1$, on obtient une application, conservant l'aire, d'une couronne ordinaire sur elle-même. L'application non perturbée est une rotation sur chaque cercle $I = \text{const}$. La condition de non-dégénérescence signifie ici que l'angle de rotation varie d'un cercle à l'autre.

Lorsque $n = 1$, les tores invariants sont des cercles ordinaires. Le théorème garantit alors que toutes les images d'un point par itération de l'application resteront au voisinage du cercle sur lequel se trouvait initialement ce point si la perturbation est suffisamment petite.

4. Voisinage d'une position d'équilibre (cas autonome). La position d'équilibre est supposée stable dans l'approximation linéaire de sorte que sont définies n fréquences propres $\omega_1, \ldots, \omega_n$. On suppose qu'entre elles il n'existe pas de relations de résonance

$$k_1\omega_1 + \ldots + k_n\omega_n = 0, \text{ les } k_i \text{ sont entiers, } 0 < \sum |k_i| \leqslant 4.$$

Le hamiltonien peut alors être ramené à la forme normale de Birkhoff (voir Appendice 7)

$$H = H_0(\tau) + \ldots,$$

où $H_0(\tau) = \sum \omega_k \tau_k + \frac{1}{2} \sum \omega_{kl} \tau_k \tau_l$, les points de suspension représentant les termes de puissance supérieure à quatre par rapport à la distance à la position d'équilibre.

La condition de non-dégénérescence

$$\det |\omega_{kl}| \neq 0$$

garantit l'existence d'un ensemble de tores invariants de mesure presque complète dans un voisinage suffisamment petit de la position d'équilibre.

La condition de non-dégénérescence isoénergétique

$$\det \begin{vmatrix} \omega_{kl} & \omega_k \\ \omega_l & 0 \end{vmatrix} \neq 0$$

garantit l'existence d'un ensemble identique de tores invariants sur chaque ensemble de niveau d'énergie (suffisamment proche d'un niveau critique).

Si $n = 2$, la condition de non-dégénérescence isoénergétique revient à exiger de la partie quadratique de H_0 qu'elle ne soit pas divisible par la linéaire. Dans ce cas la non-dégénérescence isoénergétique garantit la stabilité de la position d'équilibre au sens de Liapounov.

5. Voisinage d'une position d'équilibre (cas périodique). On suppose de nouveau qu'est réalisée la stabilité en approximation linéaire, de sorte que sont définies n fréquences propres $\omega_1, \ldots, \omega_n$. On suppose qu'entre les fréquences propres et la fréquence de variation des coefficients (que nous supposerons égale à l'unité) il n'existe de relations de résonance

$$k_1\omega_1 + \ldots + k_n\omega_n + k_0 = 0 \quad \text{avec} \quad 0 < \sum_{i=1}^{n} |k_i| \leqslant 4.$$

Le hamiltonien peut alors être ramené à la forme normale de Birkhoff qui est la même que dans le cas autonome mais avec un terme résiduel 2π-périodique en le temps.

La condition de non-dégénérescence

$$\det |\omega_{kl}| \neq 0$$

garantit l'existence de $(n + 1)$-tores invariants dans l'espace des phases élargi de dimension $(2n + 1)$, voisins du cercle $\tau = 0$ qui représente la position d'équilibre.

Lorsque $n = 1$, la condition de non-dégénérescence se traduit par la non-nullité de la dérivée de la période des petites oscillations par rapport au carré de leur amplitude. La non-dégénérescence garantit dans ce cas la stabilité de la position d'équilibre au sens de Liapounov.

6. Point fixe d'une application. On suppose ici que les $2n$ valeurs propres de l'application canonique linéarisée sont toutes de module 1 en un point fixe et ne satisfont pas à des relations de résonance d'ordre inférieur :

$$\lambda_1^{k_1} \ldots \lambda_n^{k_n} = 1, \quad |k_1| + \ldots + |k_n| \leqslant 4$$

(où les $2n$ valeurs propres sont $\lambda_1, \ldots, \lambda_n, \bar{\lambda}_1, \ldots, \bar{\lambda}_n$).

Si donc l'on néglige les puissances supérieures à trois de son développement taylorien, l'application s'écrira sous la forme normale de Birkhoff

$$(\tau, \varphi) \mapsto (\tau, \varphi + \alpha(\tau)) \text{ où } \alpha(\tau) = \partial S/\partial \tau,$$

$S = \sum \omega_k \tau_k + \frac{1}{2} \sum \omega_{kl} \tau_k \tau_l$ (les coordonnées usuelles au voisinage de la position d'équilibre sont $p_k = \sqrt{2\tau_k} \cos \varphi_k$, $q_k = \sqrt{2\tau_k} \sin \varphi_k$).

La condition de non-dégénérescence

$$\det |\omega_{kl}| \neq 0$$

garantit l'existence de n-tores invariants (voisins des tores $\tau = \text{const}$) qui forment un ensemble de mesure presque complète dans un voisinage suffisamment petit de la position d'équilibre.

Si $n = 1$, cette application est celle d'un plan ordinaire sur lui-même, et les tores invariants sont des cercles. La condition de non-dégénérescence veut dire que pour la forme normale la dérivée de l'angle de rotation d'un cercle par rapport à l'aire de ce cercle est non nulle (au point fixe et, partant, dans un de ses voisinages).

Lorsque $n = 1$, la condition de non-dégénérescence garantit la stabilité du point fixe de l'application au sens de Liapounov. Remarquons que dans ce cas la condition de non-existence des résonances d'ordre inférieur s'écrit

$$\lambda^3 \neq 1, \quad \lambda^4 \neq 1.$$

Donc tout point fixe d'une application, conservant l'aire, du plan sur lui-même est stable au sens de Liapounov si la partie linéaire de cette application est une rotation d'un angle non multiple de 90° et 120° et si est non nul le coefficient ω_{11} de la forme normale de Birkhoff (ce qui assure la dépendance non triviale de l'angle de rotation par rapport au rayon).

A aucun moment nous n'avons parlé des conditions de différentiabilité supposées dans ces théorèmes. La différentiabilité nécessaire minimale n'est connue dans aucun cas. A titre d'exemple indiquons que la dernière proposition sur la stabilité des points fixes des applications du plan sur lui-même a été d'abord démontrée par J. Moser sous l'hypothèse d'une 333-uple différentiabilité et ce n'est qu'ensuite (grâce à Moser et Russman) que le nombre de dérivées a été réduit jusqu'à 6.

E. Applications du théorème des tores invariants et de ses généralisations. Les théorèmes formulés ci-dessus s'appliquent à de nombreux problèmes de mécanique. Comme exemple considérons le mouvement d'un pendule sous l'action d'un champ extérieur périodique ou sous l'action des oscillations verticales du point d'attache.

On sait qu'en l'absence de résonance paramétrique, la position inférieure d'équilibre du pendule est stable en approximation liné-

aire. La *stabilité de cette position d'équilibre compte tenu des effets non linéaires* (à condition que n'existent pas les résonances d'ordre trois et quatre) ne peut être démontrée qu'à l'aide des théorèmes des tores invariants.

De même on peut se servir du théorème des tores invariants pour étudier les mouvements quasi périodiques d'un système d'oscillateurs non linéaires liés.

Un autre exemple est celui du flot géodésique sur une surface convexe voisine d'un ellipsoïde. Ce système est à deux degrés de liberté et nous nous assurons que *la plupart des géodésiques sur une surface voisine d'un ellipsoïde à trois axes oscillent entre deux* « caustiques » *voisines des lignes de courbure de la surface et qu'elles sont partout denses dans la couronne comprise entre ces caustiques.* Nous débouchons ainsi sur les théorèmes relatifs à la stabilité de deux géodésiques fermées déduites par déformation d'une surface composée de deux ellipses contenant l'axe moyen de l'ellipsoïde (en l'absence de résonances d'ordre 3 et 4).

Un quatrième exemple est celui des trajectoires fermées sur une table de billard de forme convexe quelconque. Parmi les trajectoires de billard fermées il en est qui sont stables en approximation linéaire et nous pouvons conclure qu'en général elles sont réellement stables. Un exemple d'une telle trajectoire stable est le petit axe d'une ellipse. Par conséquent *une trajectoire de billard fermée voisine du petit axe d'une ellipse est stable si ce billard est voisin d'une ellipse.*

L'application des théorèmes des tores invariants à la rotation d'un solide pesant asymétrique permet d'étudier le cas non intégrable d'un solide en rotation rapide. La rotation rapide équivaut mathématiquement à un mouvement de vitesse modérée dans un champ de pesanteur de faible intensité : le paramètre essentiel est le rapport de l'énergie potentielle à l'énergie cinétique. Si ce paramètre est petit, nous pouvons utiliser comme première appoximation le mouvement eulérien du solide.

En appliquant les théorèmes des tores invariants au problème à deux degrés de liberté que l'on obtient après élimination d'une coordonnée cyclique (la rotation autour de la verticale), nous obtenons la conclusion suivante quant au mouvement d'un solide en rotation rapide : *si l'énergie cinétique de rotation d'un solide est assez grande devant son énergie potentielle, alors le module du vecteur moment cinétique et sa pente restent indéfiniment au voisinage de leurs valeurs initiales.*

Il en résulte que le mouvement du solide sera indéfiniment voisin d'une combinaison de mouvements d'Euler-Poinsot à précession azimutale, excepté bien sûr le cas où les valeurs initiales de l'énergie cinétique et du moment total sont proches de celles pour lesquelles le solide est susceptible de tourner autour de son axe moyen de symétrie. Dans ce dernier cas qui n'est réalisé que pour des conditions

initiales spéciales, la bifurcation des séparatrices au voisinage de l'axe moyen engendre une culbute près de l'axe moyen plus complexe que dans le mouvement d'Euler-Poinsot.

Une généralisation du théorème des tores invariants est le théorème relatif à l'invariance adiabatique indéfinie de la variable d'action dans un système oscillatoire unidimensionnel à paramètres périodiques. On supposera ici que la loi de variation des paramètres est traduite par une fonction périodique différentiable fixée du temps lent et que le petit paramètre est le rapport de la période des oscillations propres à la période de variation des paramètres.

Alors *si la période de variation des paramètres est suffisamment grande, la variation de l'invariant adiabatique du point représentatif reste petite sur un intervalle de temps indéfiniment grand.*

On démontre de façon analogue l'*invariance adiabatique indéfinie de la variable d'action pour une particule chargée en mouvement dans un champ magnétique à symétrie axiale.* Si la symétrie axiale n'est pas réalisée, le nombre de degrés de liberté passe de deux à trois, de sorte que les tores invariants cessent de partager la variété de niveau d'énergie et l'errement du point représentatif entre les zones résonnantes devient essentiel.

Enfin, s'agissant du problème des trois (ou de plusieurs) corps, on arrive à trouver des *mouvements quasi périodiques de « type planétaire »*. Pour décrire ces mouvements il y a lieu de dire d'abord quelques mots sur l'approximation du problème du mouvement des planètes qui est plus exacte que l'approximation keplérienne. Par souci de simplicité nous nous limiterons au problème plan.

Considérons dans chaque ellipse de Kepler le vecteur qui joint le foyer (i.e. le Soleil) au centre. Ce vecteur qu'on appelle *vecteur de Laplace* caractérise aussi bien l'excentricité de l'orbite que la direction du périhélie.

L'interaction des planètes fait que toute ellipse keplérienne (et par conséquent le vecteur de Laplace) varie légèrement en fonction du temps. Il existe d'ailleurs une grande différence entre les variations du demi grand axe et celles du vecteur de Laplace. Plus exactement, le demi grand axe ne possède pas de perturbations séculaires, i.e. en première approximation il ne fait qu'osciller légèrement autour de sa valeur moyenne (« théorème de Laplace »). Le vecteur de Laplace, lui, effectue des oscillations aussi bien périodiques que séculaires. On obtient le mouvement séculaire si l'on répartit la masse de chaque planète sur son orbite proportionnellement au temps mis pour parcourir une portion d'orbite et si l'on remplace l'attraction des planètes par l'attraction des anneaux obtenus, i.e. si l'on prend les moyennes des perturbations sur les mouvements rapides. Le mouvement réel du vecteur de Laplace se déduit du mouvement séculaire par superposition des petites oscillations ; ces oscillations sont essentielles si l'intervalle de temps considéré est petit (de quelques années)

mais leur effet est petit devant celui du mouvement séculaire si l'intervalle de temps est grand (de quelques millénaires).

Les calculs (effectués encore par Lagrange) montrent que le mouvement séculaire du vecteur de Laplace de chacune des n planètes dans un même plan consiste en ce qui suit (si l'on néglige les carrés des excentricités des orbites devant ces excentricités mêmes).

Sur le plan de l'orbite de la planète il faut porter n vecteurs de module fixe, en rotation uniforme chacun avec sa vitesse angulaire. Le vecteur de Laplace est la somme de ces vecteurs.

Le vecteur de Laplace admet une telle description de son mouvement parce que le système hamiltonien qui en décrit le mouvement séculaire et dont on a pris la moyenne sur les mouvements rapides possède une position d'équilibre correspondant aux excentricités nulles. Le mouvement du vecteur de Laplace est la décomposition des petites oscillations en les oscillations propres au voisinage de la position d'équilibre mentionnée. Les vitesses angulaires et les longueurs des composantes du vecteur de Laplace (qui sont animées d'un mouvement de rotation uniforme) sont respectivement les fréquences propres et les amplitudes des oscillations propres.

Remarquons que le mouvement du vecteur de Laplace de la Terre est apparemment l'un des facteurs auquel sont liées les périodes glaciaires. En effet, lorsque l'excentricité de l'orbite de la Terre croît, le temps qu'elle passe au voisinage du Soleil diminue, et celui qu'elle passe loin du Soleil augmente (en vertu de la loi des aires) ; donc le climat devient plus rude lorsque l'excentricité augmente. La grandeur de cet effet est telle que par exemple la quantité d'énergie solaire reçue en un an à la latitude de Léningrad peut atteindre des valeurs qui correspondent actuellement aux latitudes de Kiev (si l'excentricité diminue) et de Taïmyr (si elle augmente). Le temps qui caractérise les variations de l'excentricité (des dizaines de milliers d'années) concorde bien avec les périodes glaciaires.

En vertu des théorèmes des invariants on conclut donc que, si la masse des planètes est suffisamment petite, il existe dans l'espace des phases du problème un ensemble de mesure positive, adhérence de courbes de phase quasi périodiques telles que le mouvement des planètes correspondant soit voisin du mouvement suivant des ellipses de petites excentricités et à variation lente ; quant au mouvement des vecteurs de Laplace, il est voisin du mouvement décrit par l'approximation ci-dessus. Si de surcroît les masses des planètes sont suffisamment petites, les mouvements du type décrit remplissent la plus grande partie du domaine de l'espace des phases qui, dans l'approximation keplérienne, correspond aux mouvements des planètes dans un sens sur des ellipses de petites excentricités ne se coupant pas.

Dans le problème plan des n planètes, le nombre de degrés de liberté est égal à $2n$ si l'on suppose le Soleil fixe. L'intégrale du moment cinétique permet d'éliminer une coordonnée cyclique, cependant il reste encore trop de variables pour que les tores invariants divisent la variété de niveau d'énergie (même si les planètes ne sont que deux,

cette variété est de dimension cinq et les tores de dimension trois). Donc dans le problème considéré on ne peut tirer une conclusion quant à l'invariance des demi grands axes sur un intervalle de temps infini que pour la majorité et non pas pour toutes les conditions initiales.

Le problème à deux degrés de liberté s'obtient par une idéalisation ultérieure. Remplaçons l'une des deux planètes par un «astéroïde » qui se déplacera dans le champ de l'autre planète (que nous supposerons être Jupiter) sans en perturber le mouvement.

Ce faisant nous obtenons le *problème restreint des trois corps.* Le problème nous amène à considérer pour le mouvement de l'astéroïde un système à deux degrés de liberté dépendant périodiquement du temps. Si de surcroît l'orbite de Jupiter est circulaire, alors dans des coordonnées tournant avec lui, le mouvement de l'astéroïde est décrit par un système hamiltonien autonome à deux degrés de liberté; on a donc le problème circulaire restreint plan des trois corps.

Dans ce problème le rôle du petit paramètre est tenu par le rapport de la masse de Jupiter à celle du Soleil. A la valeur nulle du paramètre correspond un mouvement keplérien non perturbé de l'astéroïde qui est représenté dans notre espace des phases de dimension quatre par un mouvement quasi périodique sur un 2-tore (puisque le système de coordonnées est en rotation). L'une des fréquences de ce mouvement quasi périodique est la même pour toutes les conditions initiales: il s'agit de la vitesse angulaire de rotation du système de coordonnées, i.e. de la fréquence de révolution de Jupiter autour du Soleil. La deuxième fréquence dépend des conditions initiales (c'est la fréquence de révolution de l'astéroïde autour du Soleil) et varie sur une variété fixée de dimension trois de niveau d'énergie de la fonction de Hamilton.

Donc la condition de non-dégénérescence n'est pas réalisée ici, en revanche la condition de non-dégénérescence isoénergétique l'est. On peut par conséquent appliquer le théorème de Kolmogorov et conclure que la plupart des tores invariants à rapport de fréquences irrationnel se conservent lorsque la masse de la planète perturbatrice (Jupiter) est différente de zéro, mais suffisamment petite.

Par ailleurs les 2-tores invariants divisent la variété de dimension trois de niveau d'énergie du hamiltonien. Donc le *demi grand axe et l'excentricité de l'ellipse keplérienne de l'astéroïde resteront indéfiniment au voisinage de leurs valeurs initiales si, à la date initiale, l'ellipse keplérienne n'a pas coupé l'orbite de la planète perturbatrice et si la masse de cette dernière est suffisamment petite.*

L'ellipse keplérienne de l'astéroïde est susceptible de tourner lentement dans un système de coordonnées fixe, puisque notre système n'est qu'isoénergétiquement non dégénéré et par conséquent une perturbation du tore invariant conserve non pas les fréquences mais leur rapport. Sous l'effet de cette perturbation la fréquence du mouvement azimutal du périhélie de l'astéroïde peut, dans un système de coordonnées mobile, être légèrement différente de la fréquence de Jupiter et à ce moment le périhélie tournera lentement dans un système fixe.

APPENDICE 9

THÉORÈME GÉOMÉTRIQUE DE POINCARÉ, GÉNÉRALISATION ET APPLICATIONS

L'étude des solutions périodiques des problèmes de mécanique céleste a conduit H. Poincaré à élaborer un modèle relativement simple qui contenait déjà la difficulté principale. Ce modèle est une application, conservant l'aire, d'une couronne plane sur elle-même.

Les applications de ce type s'introduisent lorsqu'on étudie les systèmes dynamiques à deux degrés de liberté. Plus exactement, l'application d'une 2-surface de section sur elle-même se construit de la manière suivante: chaque point de la surface de section se transforme en le point d'intersection suivant de la courbe de phase issue du premier point, avec la surface de section (voir Appendice 7).

Ceci étant, aux courbes de phase fermées correspondent les points fixes d'une application ou de ses puissances. Et inversement, tout point fixe d'une application ou de ses puissances définit une courbe de phase fermée.

Donc l'étude de l'existence de solutions périodiques des problèmes dynamiques se ramène à l'étude des points fixes d'applications, conservant l'aire, d'une couronne sur elle-même.

L'étude de ces applications a conduit Poincaré à formuler le théorème suivant.

A. Points fixes de l'application d'une couronne sur elle-même.

Théorème. *Soit donnée une application homéomorphe, conservant l'aire d'une couronne plane sur elle-même. Supposons que les bords de cette couronne tournent l'un en avant, l'autre en arrière par cette application. Alors cette application possède au moins deux points fixes.*

L'hypothèse que les bords de la couronne tournent l'un en avant et l'autre en arrière signifie que si l'on choisit les coordonnées (x, y mod 2π) telles que ces bords aient pour équations $x = a$ et $x = b$, alors cette application sera définie par les formules

$$(x, y) \mapsto (f(x, y), y + g(x, y)),$$

où les fonctions f et g sont continues 2π-périodiques en y, et telles que $f(a, y) \equiv a$, $f(b, y) \equiv b$ et $g(a, y) < 0$, $g(b, y) > 0$ pour tous les y.

La démonstration de ce théorème que Poincaré a publié un peu avant sa mort n'a été faite que plus tard par G. D. Birkhoff *).

Nombre de questions soulevées par ce théorème sont encore en suspens et notamment sa généralisation qui est très importante pour l'étude des solutions périodiques des problèmes ayant un grand nombre de degrés de liberté.

Le fait est que les raisonnements qui ont amené Poincaré à ce théorème s'appliquent à une foule d'autres cas, alors que la démonstration ingénieuse proposée par Birkhoff se généralise mal. C'est pourquoi on ignore si les déductions suggérées par le raisonnement de Poincaré sont valables en dehors du cadre du théorème relatif à la couronne plane. Voici ce raisonnement.

B. Lien entre les points fixes d'une application et les points critiques de sa fonction génératrice. Définissons un difféomorphisme symplectique de la couronne

$$(x, y) \mapsto (X, Y)$$

par la fonction génératrice $Xy + S(X, y)$, où la fonction S est 2π-périodique en y. Pour cette écriture du difféomorphisme il faut que $\partial X/\partial x \neq 0$. Donc

$$dS = (x - X)\, dy + (Y - y)\, dX,$$

et par conséquent les points fixes du difféomorphisme sont les points critiques de la fonction $F(x, y) = S(X(x, y), y)$. On peut toujours construire la fonction F en la considérant comme l'intégrale de la forme $(x - X)\, dy + (Y - y)\, dX$. Le gradient de cette fonction est dirigé soit vers l'intérieur de la couronne soit vers l'extérieur simultanément sur les deux cercles de la couronne (en vertu de la condition de rotation dans des sens opposés).

Or toute fonction différentiable dans une couronne, dont le gradient sur les bords est dirigé vers l'intérieur (ou vers l'extérieur), possède un point critique (un minimum ou un maximum) à l'intérieur. Bien plus, on démontre que dans la couronne une telle fonction possède au moins deux points critiques. Donc on aurait pu affirmer que notre difféomorphisme possède au moins deux points fixes si l'on avait été sûr que chaque point critique de la fonction F est un point fixe de l'application.

Ceci n'est malheureusement vrai qu'à la condition supplémentaire que $\partial X/\partial x \neq 0$, condition qui nous permet d'exprimer F en fonction de X et y. Donc notre raisonnement est valable pour les

*) G. D. Birkhoff, *Dynamical systems*, New York (1927).

applications qui ne s'écartent pas trop de l'identité. Il suffit par exemple que les dérivées de la fonction génératrice F soient inférieures à 1.

Une amélioration du même raisonnement (avec un autre choix de la fonction génératrice *)) montre qu'il suffit que les valeurs propres du jacobien $D(X, Y)/D(x, y)$ ne soient en aucun point égales à -1, i.e. que notre application ne fasse tourner l'espace tangent en aucun point. Malheureusement toutes ces conditions ne sont pas partout réalisées pour les applications qui ne sont pas voisines de l'identité. La démonstration du théorème de Poincaré dans le cas général fait appel à des raisonnements totalement différents.

Le lien existant entre les points fixes d'une application et les points critiques des fonctions génératrices semble en soi plus important que le théorème sur les applications d'une couronne plane.

Plus bas on cite quelques exemples dans lesquels ce lien conduit à d'importantes conclusions, à vrai dire sous certaines restrictions dont la nécessité n'est pas très claire.

C. **Difféomorphismes symplectiques du tore.** Considérons un difféomorphisme symplectique du tore laissant fixe le centre de gravité

$$(x, y) \longmapsto (x + f(x, y)), y + g(x, y) = (X, Y),$$

où x et y mod 2π sont des coordonnées angulaires sur le tore, la symplecticité signifie que $D(X, Y)/D(x, y) = 1$ et la condition d'invariance du centre de gravité entraîne que les valeurs moyennes des fonctions f et g sont nulles.

Théorème. *Un tel difféomorphisme possède au minimum quatre points fixes, comptés avec leurs multiplicités dont au moins trois ne sont pas confondus pourvu que les valeurs propres du jacobien ne soient égales à -1 en aucun point.*

La démonstration se base sur l'étude d'une fonction définie sur le tore par la formule

$$\Phi(x, y) = \frac{1}{2} \int (X - x)(dY + dy) - (Y - y)(dX + dx),$$

et sur le fait qu'une fonction différentiable sur un tore possède au minimum quatre points critiques (comptés avec leurs multiplicités) dont au moins trois ne sont pas confondus.

La démonstration de ce théorème sans restrictions sur les valeurs propres pose les mêmes difficultés que celles auxquelles a été confronté Poincaré dans le théorème de la couronne.

*) $d\Phi = \frac{1}{2} \begin{vmatrix} X - x & Y - y \\ dX + dx & dY + dy \end{vmatrix}.$

Remarquons que le théorème de la couronne découlerait du théorème du tore si, dans ce dernier, il était possible de s'affranchir de la condition sur les valeurs propres. En effet, composons un tore avec deux exemplaires de notre couronne en disposant au voisinage de chacun des deux cercles frontières une étroite couronne joint. Nous pouvons alors prolonger notre application de la couronne jusqu'à un difféomorphisme du tore tel que: 1) sur chacune des deux grandes couronnes le difféomorphisme coïncide avec l'initial, 2) sur chacune des deux couronnes joints le difféomorphisme ne possède pas de points fixes, 3) le centre de gravité reste en place.

La construction d'un tel difféomorphisme du tore utilise la propriété des bords de la couronne de tourner dans des sens opposés. Sur chaque couronne joint, les points se déplacent dans le même sens que sur ses bords. Comme sur les deux couronnes joints le mouvement a lieu dans des sens opposés, on peut le choisir de manière à assurer l'invariance du centre de gravité.

A présent, sur les quatre points fixes du tore, deux doivent se trouver sur la couronne initiale, et du théorème sur le tore nous déduisons le théorème sur la couronne.

Le théorème relatif au tore se généralise à d'autres variétés symplectiques bidimensionnelles et multidimensionnelles. Pour énoncer ces généralisations il faut commencer par reformuler la condition d'invariance du centre de gravité.

Soit $g: M \to M$ un difféomorphisme symplectique. Nous dirons que le difféomorphisme g est *homologue à l'identité* si on peut le joindre à l'identité (laissant fixes tous les points de la variété M) par une courbe différentiable g_t composée de difféomorphismes symplectiques de telle sorte que le champ de vitesses $\dot{g}_t$ possède à chaque instant t un hamiltonien univoque. On démontre que les difféomorphismes symplectiques homologues à l'identité forment un commutant de la composante connexe de l'unité dans le groupe des difféomorphismes symplectiques de la variété M.

Si la variété considérée est un 2-tore, les difféomorphismes homologues à l'identité sont exactement ceux dont on a dit plus haut qu'ils conservaient le centre de gravité.

Nous sommes arrivés ainsi à la généralisation suivante du théorème de Poincaré.

Théorème. *Tout difféomorphisme symplectique d'une variété symplectique compacte, homologue à l'identité, possède au moins autant de points fixes qu'une fonction différentiable sur cette variété possède de points critiques, pourvu qu'il ne soit pas trop éloigné de l'identité.*

Remarquons que la conduction d'homologie à l'identité est essentielle comme cela découle du déplacement ne laissant invariant aucun point sur le tore.

Pour ce qui est de la dernière condition (le difféomorphisme n'est pas trop éloigné de l'identité), on ne sait pas si elle est essentielle ou non. Lorsque la variété envisagée est un $2n$-tore, il suffit qu'aucune des valeurs propres du jacobien du difféomorphisme (dans un quelconque système de coordonnées symplectique global défini sur $\mathbf{R}^{2n}$) ne soit égale à -1.

Une telle restriction est sans doute nécessaire dans les problèmes multidimensionnels, car il n'est pas exclu que le théorème de Poincaré soit essentiellement bidimensionnel tout comme le théorème suivant de A. Chnirelman et N. Nikichine. *Tout difféomorphisme, conservant l'aire, d'une 2-sphère sur elle-même possède au minimum deux points fixes non confondus.*

La démonstration de ce théorème se base sur le fait que l'indice du champ de vecteurs du gradient d'une fonction différentiable de deux variables en un point critique isolé ne peut être supérieur à l'unité (quoiqu'il puisse prendre les valeurs 1, 0, —1, —2, —3, ...) et la somme des indices de tous les points fixes d'un difféomorphisme, conservant l'orientation, d'une 2-sphère sur elle-même est égale à deux.

L'indice du gradient d'une fonction différentiable d'un nombre plus grand de variables peut prendre des valeurs entières quelconques en un point critique.

D. Intersections des variétés lagrangiennes. On peut donner une autre forme au raisonnement de Poincaré si l'on considère les points de la couronne qui se déplacent exclusivement radialement. De tels points existent sur chaque rayon, puisque les cercles frontières tournent dans des sens opposés. Supposons que nous ayons réussi à construire avec les points se déplaçant radialement une courbe fermée séparant les deux cercles frontières de la couronne. L'image de cette courbe par notre application doit couper cette courbe (puisque les domaines en lesquels elle partage la couronne se transforment en des domaines d'aires égales).

Si la courbe indiquée et son image coupent chaque rayon une seule fois, les points d'intersection de la courbe avec cette image sont de toute évidence les points fixes de l'application.

Certains éléments de ce raisonnement peuvent être généralisés, ce qui fournit des résultats intéressants sur les solutions périodiques des problèmes de dynamique. Dans le cas multidimensionnel la couronne est un espace des phases, produit cartésien d'un domaine d'un espace euclidien par un tore de dimension commune (la couronne est le produit d'un intervalle par un cercle). La structure symplectique sur l'espace des phases est définie comme d'habitude, i.e. est de la forme $\Omega = \Sigma dx_k \wedge dy_k$, où x_k sont des variables d'action et Y_k des variables angulaires.

Il n'est pas difficile d'établir lesquels des difféomorphismes symplectiques de notre espace des phases sont homologues à l'identité. Plus exactement, un difféomorphisme symplectique A est homologue à l'identité si on peut le déduire de l'identité par une déformation continue et si de plus

$$\oint_{\gamma} x\, dy = \oint_{A\gamma} x\, dy$$

pour tout contour fermé (pas nécessairement homologue à zéro).

La condition d'homologie à l'identité interdit tout déplacement systématique le long d'une x-direction (l'« évolution des variables d'action ») et n'autorise que les déplacements sur les tores.

Considérons un n-tore quelconque $x = c =$ const et appliquons-lui notre difféomorphisme symplectique homologue à l'identité. On obtient de nouveau un n-tore. Il s'avère que *le tore initial coupe son image au moins en 2^n points (comptés avec leurs multiplicités) dont $(n + 1)$ au moins ne sont pas confondus, pourvu que l'équation du tore image soit de la forme $x = f(y)$, où f est différentiable.*

Pour $n = 1$ cela veut dire que chacun des cercles frontières de la couronne coupe son image au moins en deux points. Ceci découle immédiatement de la conservation de l'aire, la condition que l'équation de l'image soit de la forme $x = f(y)$ n'étant plus nécessaire.

Cette condition est-elle indispensable dans le cas multidimensionnel? Il est difficile de se prononcer. Cependant, si on l'envisage, la démonstration se déroule comme suit.

Remarquons que le tore initial est une sous-variété lagrangienne de l'espace des phases. Notre difféomorphisme étant symplectique, le tore image est également lagrangien et la 1-forme $(x - c)\, dy$ définie sur lui est fermée. Bien plus, c'est la différentielle totale d'une fonction différentiable *univoque* F de sorte que notre difféomorphisme est homologue à l'identité et pour tout contour fermé on a donc

$$\oint_{A\gamma} (x-c)\, dy = \oint_{A\gamma} x\, dy - \oint_{A\gamma} c\, dy = \oint_{\gamma} x\, dy - \oint_{A\gamma} c\, dy = c \oint_{\gamma} dy - c \oint_{A\gamma} dy = 0.$$

Remarquons que les points d'intersection du tore avec son image sont les points critiques de la fonction F (puisqu'en ces points $dF = (x - c)dy = 0$).

De la condition d'univocité de la projection du tore image (i.e. du fait que le tore image a pour équation $x = f(y)$) il découle qu'inversement tous les points critiques de la fonction F sont les points d'intersection de nos tores. En effet, à la condition indiquée y peut être pris pour coordonnée locale sur le tore et par conséquent l'égalité $dF = 0$ pour tous les vecteurs tangents au tore image entraîne $x = c$.

Une fonction différentiable sur un n-tore possède au moins 2^n points critiques, comptés avec leurs multiplicités, dont $(n + 1)$ au moins ne sont pas confondus (voir par exemple J. Milnor, *Morse Theory*, Princeton, 1963).

Donc nos tores se coupent au moins en 2^n points (points multiples compris) dont au moins $(n + 1)$ ne sont pas confondus.

Un raisonnement analogue montre que *tout tore lagrangien est coupé par son image au moins en 2^n points (dont au moins $(n + 1)$ sont distincts), pourvu que le tore initial et son image se projettent univoquement sur un y-espace, i.e. soient définis respectivement par les équations $x = f(y)$, $x = g(y)$.*

Au reste cette proposition se ramène à la précédente par l'application canonique $(x, y) \mapsto (x - f(y), y)$.

E. Applications à la recherche des points fixes et des solutions périodiques. Considérons maintenant une transformation symplectique homologue à l'identité du même type spécial que celui qui s'introduit dans les problèmes intégrables de la dynamique, i.e. de la forme

$$A_0(x, y) = (x, y + \omega(x)), \text{ où } \omega = \partial S/\partial x.$$

Ici $x \in \mathbf{R}^n$ est une variable d'action, $y \bmod 2\pi \in T^n$ une variable angulaire.

Supposons que sur le tore $x = x_0$ toutes les fréquences sont commensurables :

$$\omega_i(x_0) = \frac{k_i}{N} 2\pi \text{ avec } k_i \text{ et } N \text{ entiers ;}$$
$$\omega(x_0) \neq 0,$$

et qu'est réalisée la condition de non-dégénérescence

$$\det |\partial\omega/\partial x|_{x_0} \neq 0.$$

Théorème. *Tout difféomorphisme symplectique A homologue à l'identité et suffisamment proche de A_0 possède au voisinage du tore $x = x_0$ au moins 2^n points périodiques ξ, points multiples compris, de période N (de sorte que $A^N\xi = \xi$).*

On pourrait ramener la démonstration à l'étude de l'intersection de deux sous-variétés lagrangiennes de l'espace de dimension $4n$ ($\mathbf{R}^n \times T^n \times \mathbf{R}^n \times T^n$) muni de la forme $\Omega = dx \wedge dy - dX \wedge dY$, dont l'une est la diagonale ($X = x$, $Y = y$) et l'autre, le graphe de l'application A^N.

Mais il est plus facile de construire directement une fonction convenable sur le tore. En effet, l'application A_0^N s'écrit

$$(x, y) \longmapsto (x, y + \alpha(x)), \text{ où } \alpha(x_0) = 0, \ \det |\partial\alpha/\partial x|_{x_0} \neq 0.$$

En vertu du théorème des fonctions implicites, l'application A^N possède au voisinage du tore $x = x_0$ un tore qui se déplace exclusivement radialement ($(x, y) \longmapsto (X, Y)$), dont l'équation est de la forme $x = f(y)$ et dont l'image ai également une équation de la même forme : $x = g(y)$. Dans ces notations $X(f(y), y) = g(y)$, $Y(f(y), y) = y$.

L'application A étant homologue à l'identité, il résulte que A^n possède une fonction génératrice globale univoque de la forme $Xy + S(X, y)$, où S est 2π-périodique en y.

La fonction $F(y) = S(X(f(y), y), y)$ possède sur le tore au minimum 2^n points critiques y_k. Tous les points $\xi_k = (f(y_k), y_k)$ sont les points fixes de A^N. En effet,

$$dF = (x - X)\,dy + (Y - y)\,dX = (x - X)\,dy = (f(y) - g(y))\,dy.$$

Donc de $dF|y_k = 0$ il s'ensuit que $f(y_k) = g(y_k)$, i.e. $A^N\xi_k = \xi_k$, c.q.f.d.

Considérons maintenant les trajectoires fermées de systèmes conservatifs. Ce résultat se formule comme suit dans les termes de l'Appendice 8.

Corollaire. *Lorsqu'un n-tore invariant, composé des trajectoires fermées d'un système isoénergétiquement non-dégénéré à n degrés de liberté, se décompose, il se forme au minimum 2^{n-1} trajectoires fermées*

du problème perturbé (comptées avec leurs multiplicités) dont au minimum n sont distinctes, pourvu que la perturbation soit suffisamment petite.

La démonstration se ramène au théorème précédent moyennant une $(2n - 2)$-surface de section. Ceci étant, il importe de choisir au départ les coordonnées angulaires y telles que les trajectoires fermées du problème non perturbé soient données sur le tore par les équations $\dot{y}_2 = \ldots = \dot{y}_n = 0$ et de déterminer la surface de section par l'équation $y_1 = 0$.

Dans le cas de deux degrés de liberté on peut appliquer le théorème de Poincaré aux couronnes engendrées par l'intersection des invariants avec une 2-surface. Nous obtenons le résultat suivant.

Dans la faille comprise entre deux tores invariants de dimension deux d'un système à deux degrés de liberté, il existe au minimum deux trajectoires de phase fermées si les rapports des fréquences des mouvements quasi périodiques ne sont pas égaux sur ces tores.

Ainsi nous obtenons d'innombrables solutions périodiques dans tous les problèmes à deux degrés de liberté où sont connus les tores invariants (par exemple, dans le problème restreint circulaire des trois corps, le problème des géodésiques fermées, etc.). On avance même l'hypothèse que les courbes de phase fermées des systèmes hamiltoniens génériques dont l'espace des phases est compact forment un ensemble partout dense. D'ailleurs, s'il en est ainsi, la fermeture de la plupart de ces courbes n'est pas essentielle puisque leurs périodes sont extrêmement grands.

La méthode de Poincaré s'applique aux systèmes à nombre de degrés de liberté plus grand que deux: exemple, le théorème de Birkhoff sur l'existence d'une infinité de solutions voisines d'une solution périodique linéaire stable générique (ou sur l'existence d'une infinité de points périodiques au voisinage d'un point fixe d'une application symplectique non dégénéré linéaire et stable de l'espace sur lui-même). La démonstration consiste à approximiser d'abord l'application par sa forme normale et à utiliser ensuite le lien entre les points fixes de cette application et les points critiques de sa fonction génératrice.

La connaissance des solutions périodiques permet entre autres de démontrer la non-existence des intégrales premières (différentes des classiques) dans de nombreux problèmes de dynamique. Supposons par exemple que sur une variété quelconque de niveau des intégrales connues l'on ait découvert une trajectoire périodique instable. Dans le cas général ses séparatrices forment un réseau compliqué que nous avons étudié dans l'Appendice 7. Si l'on découvre le phénomène de bifurcation des séparatrices et si l'on démontre que ces séparatrices ne tiennent dans aucune variété de dimension inférieure à celle de la variété considérée, alors on peut être certain que le système ne possède pas d'autres intégrales premières.

Au reste, le comportement compliqué des courbes de phase qui interdit l'existence des intégrales premières est souvent facile à mettre en évidence sans l'aide des solutions périodiques, simplement par un regard sur le croquis, établi sur ordinateur, de l'intersection d'une courbe de phase avec une surface de section.

F. Invariance de la fonction génératrice. Nous avons déjà souligné le désagrément dû à non-invariance des fonctions génératrices relativement au choix des coordonnées canoniques sur une variété symplectique.

D'un autre côté, nous avons à maintes reprises utilisé le lien existant entre les points fixes d'une application et les points critiques de sa fonction génératrice.

Cependant, même si d'une façon générale la fonction génératrice n'est pas intrinsèquement liée à l'application, il s'avère qu'il en va autrement au voisinage d'un point fixe.

Plus exactement, soit donné un difféomorphisme symplectique laissant fixe un certain point. Définissons au voisinage de ce point la « fonction génératrice »

$$\Phi = \frac{1}{2} \int \sum \begin{vmatrix} X_k - x_k & Y_k - y_k \\ dX_k + dx_k & dY_k + dy_k \end{vmatrix}$$

à l'aide d'un certain système symplectique de coordonnées (x, y) *).

Construisons ensuite d'une façon analogue une autre fonction génératrice Φ' à l'aide d'un autre système symplectique de coordonnées (x', y').

Théorème. *Si le linéarisé d'un difféomorphisme symplectique ne possède pas en un point fixe de valeurs propres égales à* -1, *alors les fonctions* Φ *et* Φ' *sont équivalentes dans son voisinage en ce sens qu'is existe un difféomorphisme* g (*en général non symplectique*) *tel qul*

$$\Phi(z) \equiv \Phi'(g(z)) + \text{const.}$$

La démonstration figure dans l'article de A. Weinstein, *The invariance of Poincaré's generating function for canonical transformations.* Inventiones Mathematicae, **16**, n° 3 (1972), 202-214.

A remarquer que deux difféomorphismes dont les fonctions génératrices sont équivalentes au voisinage d'un point fixe ne sont pas forcement équivalents dans la classe des difféomorphismes symplectiques (exemple : la rotation et la rotation d'un angle dépendant du rayon, avec les parties quadratiques de la fonction génératrice non dégénérées en 0).

---

*) L'accroissement de cette fonction le long d'un arc quelconque est égal à l'intégrale d'une 1-forme, définissant une structure symplectique, étendue à une bande formée par les segments de droite qui joignent tout point à son image. Donc une telle fonction Φ est intrinsèquement liée à l'application relativement aux changements linéaires des coordonnées canoniques.

APPENDICE 10

MULTIPLICITÉ DES FRÉQUENCES PROPRES ET DES ELLIPSOÏDES DÉPENDANT DES PARAMÈTRES

Nous avons évoqué à maintes reprises les familles d'ellipsoïdes dans l'espace euclidien. En étudiant la dépendance des fréquences propres des petites oscillations par rapport aux paramètres nous avons été confrontés à l'ellipsoïde de niveau d'énergie potentielle qui dépend de la rigidité du système dans l'espace euclidien (la métrique de l'espace est définie par l'énergie cinétique). Un autre exemple est l'ellipsoïde d'inertie du solide (les paramètres sont ici la forme du solide et la distribution des masses sur lui).

Penchons-nous maintenant sur le problème général suivant: pour quelles valeurs des paramètres le spectre des valeurs propres dégénère-t-il, i.e. pour quelles valeurs l'ellipsoïde correspondant se transforme-t-il en ellipsoïde de révolution? Remarquons que les valeurs propres d'une forme quadratique définie sur l'espace euclidien (ou les longueurs des axes d'un ellipsoïde) varient continûment lorsque les paramètres du système (les coefficients de la forme) varient continûment. Il semble naturel de s'attendre à ce que, dans un système dépendant d'un seul paramètre, lorsque ce dernier varie, l'une des valeurs propres heurtera une autre à certains moments, de sorte que le système aura un spectre multiple pour certaines valeurs du paramètre.

Imaginons-nous par exemple que nous voulons transformer l'ellipsoïde d'inertie d'un solide en ellipsoïde de révolution en faisant coulisser une masse ajustée sur une tige rigidement fixée au solide de sorte que nous disposons d'un paramètre. Les trois principaux axes d'inertie a, b, c seront des fonctions continues de ce paramètre et de prime abord il semble qu'en choisissant convenablement le paramètre (p) on peut réaliser l'égalité de deux axes, par exemple $a(p) = b(p)$.

La réalité est cependant tout autre et en général il faut déplacer *au minimum deux masses* ajustées pour transformer l'ellipsoïde d'inertie en ellipsoïde de révolution.

D'une façon générale, *dans les familles génériques de formes quadratiques le spectre n'est multiple que pour deux paramètres ou plus, alors*

28—01408

que dans les familles génériques à un paramètre il est simple pour toutes les valeurs du paramètre. En pratique cela se traduit par le fait que lorsque le paramètre d'une famille générique à un paramètre varie, les valeurs propres peuvent se rapprocher étroitement, mais arrivées suffisamment près l'une de l'autre, elles commencent à se repousser et s'éloignent de nouveau trompant celui qui espérait obtenir un spectre multiple en faisant varier le paramètre.

Dans le présent Appendice on étudiera les causes de ce comportement, étrange à prime abord, des valeurs propres, puis on traitera brièvement de questions analogues pour des systèmes à divers groupes de symétries.

A. Variété d'ellipsoïdes de révolution. Considérons l'ensemble de toutes les formes quadratiques de l'espace euclidien $\mathbf{R}^n$. Cet ensemble possède une structure naturelle d'espace vectoriel de dimension $n(n+1)/2$. Toutes les formes quadratiques du plan constituent par exemple un espace de dimension trois (la forme $Ax^2 + 2Bxy + Cy^2$ a pour coordonnées les nombres A, B, C).

Les formes définies positives constituent un ouvert dans cet espace (dans le cas du plan c'est l'intérieur de l'une des nappes du cône $B^2 = AC$ engendré par les formes dégénérées).

Chaque ellipsoïde de centre en l'origine des coordonnées détermine une forme quadratique définie positive dont il est ensemble de niveau 1 ; réciproquement l'ensemble de niveau 1 de toute forme quadratique définie positive est un ellipsoïde. Nous pouvons donc identifier les ensembles des formes quadratiques définies positives aux ellipsoïdes centrés en l'origine des coordonnées.

Ce faisant, nous introduisons sur l'ensemble des ellipsoïdes de $\mathbf{R}^n$ centrés en O une structure de variété différentiable de dimension $n(n+1)/2$ (cette variété est recouverte par une seule carte : l'ouvert indiqué plus haut dans l'espace des formes quadratiques).

Considérons maintenant l'ensemble de tous les *ellipsoïdes de révolution.* Nous affirmons que cet ensemble est de codimension 2 dans l'espace envisagé, i.e. il est défini par deux équations indépendantes et non pas par une seule comme on pourrait le penser à première vue. Plus exactement, est vrai le

Théorème 1. *L'ensemble des ellipsoïdes de révolution est une union finie de sous-variétés différentiables de codimension* $\geqslant 2$ *dans la variété de tous les ellipsoïdes.*

Par *codimension* on entend ici la différence entre la dimension de l'espace des ellipsoïdes et la dimension d'une sous-variété.

Démonstration. Considérons tout d'abord les ellipsoïdes de $\mathbf{R}^n$ dont deux axes sont égaux et les autres distincts. Un tel ellipsoïde est défini par les directions des axes distincts, ce qui donne

$$(n-1) + (n-2) + \ldots + 2 = (n+1)(n-2)/2$$

paramètres, et par les longueurs des axes, soit $(n - 1)$ paramètres. Au total on a

$$(n^2 - n - 2 + 2n - 2)/2,$$

ce qui est de deux inférieur à la dimension $n\,(n + 1)/2$ de l'espace de tous les ellipsoïdes. La détermination du nombre de paramètres montre également que l'ensemble des ellipsoïdes dont exactement deux axes sont égaux est une variété.

S'agissant des ellipsoïdes ayant un nombre plus grand d'axes égaux, il est clair qu'ils forment tous un ensemble de dimension encore plus petite. La démonstration rigoureuse découle du lemme suivant.

Lemme. *L'ensemble de tous les ellipsoïdes possédant ν_2 axes doubles, ν_3 axes triples, ν_4 axes quadruples, etc., est une sous-variété différentiable de la variété des ellipsoïdes de codimension*

$$2\nu_2 + 5\nu_3 + 9\nu_4 + \ldots = \sum \frac{1}{2}\,(i-1)\,(i+2)\,\nu_i.$$

On démontre ce lemme en déterminant le nombre de paramètres comme dans le cas particulier précédent (qui correspond à $\nu_2 = 1$, $\nu_3 = \nu_4 = \ldots = 0$). Le lecteur effectuera lui-même ce calcul après avoir remarqué que la dimension de la variété de tous les sous-espaces de dimension k d'un espace vectoriel de dimension n est égale à $k\,(n - k)$ (puisqu'un k-plan d'un espace de dimension n peut être considéré comme le graphe d'une application d'un espace de dimension k dans un espace de dimension $(n - k)$, or cette application est définie par une $k \times (n - k)$-matrice)).

Exemple. Considérons le cas $n = 2$, i.e. des ellipses sur le plan. L'ellipse est définie par trois paramètres (par la longueur des deux axes et un angle définissant la direction de l'un d'eux). Donc, la variété des ellipses sur le plan est de dimension trois comme le laissait prévoir notre formule.

Le cercle, lui, est défini par un paramètre (le rayon). Donc la variété des cercles dans l'espace des ellipsoïdes est une ligne dans un espace de dimension trois et non pas une surface comme on pourrait le croire à première vue.

Ce paradoxe, si paradoxe il y a, devient clair à la lumière des calculs suivants. Les formes quadratiques $Ax^2 + 2Bxy + Cy^2$ dont les valeurs propres sont égales constituent dans un espace de dimension trois muni des coordonnées A, B, C une variété définie par *une seule* équation $\lambda_1 = \lambda_2$, où $\lambda_{1,2}\,(A, B, C)$ sont les valeurs propres. Cependant, le premier membre de cette équation est la somme de deux carrés comme il ressort de la formule du discriminant de l'équation caractéristique

$$\Delta = (A + C)^2 - 4\,(AC - B^2) = (A - C)^2 + 4B^2.$$

Donc la seule équation $\Delta = 0$ définit dans l'espace de dimension trois des formes quadratiques une droite ($A = C$ $B = 0$) et non une surface.

La conclusion qui découle immédiatement de ce que la variété des ellipsoïdes est de codimension 2 est que cette variété *ne divise pas* l'espace de tous les ellipsoïdes comme la droite ne divise pas l'espace de dimension trois (quant à la variété des formes à spectre multiple elle ne divise pas l'espace des formes quadratiques).

Nous pouvons donc affirmer que non seulement l'ellipsoïde générique possède des axes de longueur différente, mais encore que *deux tels ellipsoïdes peuvent être reliés dans l'espace des ellipsoïdes par une courbe différentiable composée entièrement d'ellipsoïdes ayant des axes de longueur différente.* Bien plus, si deux ellipsoïdes génériques sont reliés dans l'espace des ellipsoïdes par une courbe différentiable contenant des points qui sont des ellipsoïdes de révolution, alors, en faisant subir à cette courbe un déplacement aussi petit que l'on voudra, on peut la séparer de l'ensemble des ellipsoïdes de révolution de sorte que tous les points de la nouvelle courbe seront des ellipsoïdes sans axes multiples.

De ce qui précède découle en particulier une démonstration simple du théorème relatif à l'accroissement des fréquences propres lorsque la rigidité d'un système croît. En effet, les dérivées par rapport au paramètre d'une valeur propre simple d'une forme quadratique sont déterminées par la dérivée de la forme quadratique suivant la direction propre correspondante. Si la rigidité croît, alors l'énergie potentielle croît suivant chaque direction, et partant suivant la direction propre. Donc la fréquence propre augmente. Ainsi sans recourir au spectre multiple nous avons démontré le théorème sur l'accroissement des fréquences quand on passe du système initial à un système plus rigide. Dans le cas d'un spectre multiple, la démonstration se fait par un passage à la limite, légitime grâce au fait que l'intérieur du chemin qui conduit du système initial au système plus rigide peut être séparé de l'ensemble des systèmes à spectre multiple par un déplacement aussi petit que l'on veut.

En résumé nous pouvons conclure qu'*une famille générique à un paramètre d'ellipsoïdes* (*resp. de formes quadratiques dans l'espace euclidien*) *ne contient pas d'ellipsoïdes de révolution* (*resp. de formes quadratiques à spectre multiple*). Si l'on transpose ceci à l'ellipsoïde d'inertie on retrouve la conclusion précédente, à savoir qu'il faut déplacer deux masses ajustées pour obtenir un ellipsoïde de révolution.

Voyons maintenant les familles à deux paramètres. Des calculs effectués il découle que *dans les familles génériques à deux paramètres les ellipsoïdes de révolution ne se rencontrent qu'en des points isolés du plan des paramètres.*

Soit une surface convexe dans l'espace ordinaire. La deuxième forme quadratique de la surface définit une ellipse dans l'espace tangent en chaque point. Il s'introduit donc une famille à deux paramètres d'ellipses (que l'on peut transporter dans un plan en choisissant un système de coordonnées local au voisinage d'un point de la surface). Nous concluons donc qu'en chaque point de la

surface, à l'exception de certains points isolés, l'ellipse possède des axes de longueur différente. Donc sur la surface générique considérée sont définis deux champs orthogonaux de directions (les grands et les petits axes de l'ellipse) *à points singuliers isolés.* En géométrie différentielle ces directions sont appelées directions des courbures principales et ces points singuliers, les *points ombilicaux.* Sur l'ellipsoïde par exemple il existe quatre points ombilicaux; ils sont situés sur l'ellipse qui contient le petit et le grand axe; deux d'entre eux sont bien visibles sur la figure 207 qui représente les géodésiques sur un ellipsoïde.

D'une façon analogue, *dans une famille générique à trois paramètres, les ellipsoïdes de révolution ne se rencontrent que sur certaines lignes de l'espace de dimension trois des paramètres.* Si par exemple en chaque point de l'espace euclidien ordinaire est donné un ellipsoïde (i.e. est défini un tenseur symétrique d'ordre 2), alors les singularités des champs d'axes principaux seront, d'une façon générale, situées sur des lignes distinctes (où deux des trois champs de directions présentent une discontinuité).

Ces lignes, comme les points ombilicaux de l'exemple précédent, sont de plusieurs types. On déduit leur classification (pour les champs génériques d'ellipsoïdes) de la classification des singularités des projections lagrangiennes qui figure dans l'Appendice 12.

Dans une famille générique *à quatre paramètres* les ellipsoïdes de révolution se rencontrent sur les 2-surfaces de l'espace des paramètres. Ces surfaces ne présentent pas de singularités excepté les intersections transversales dans certains points de l'espace des paramètres; à ces valeurs des paramètres correspondent des ellipsoïdes ayant deux couples (distincts) d'axes égaux.

L'axe triple apparaît pour la première fois pour cinq paramètres, dans des points isolés de l'espace des paramètres. Les valeurs de ces paramètres qui correspondent aux ellipsoïdes à axe double forment dans l'espace de dimension cinq des paramètres une variété de dimension trois à singularités de deux espèces: des intersections transversales de deux branches le long d'une courbe et des singularités coniques dans certains points (non situés sur cette courbe), plus exactement, dans les points de l'espace des paramètres qui correspondent aux ellipsoïdes à axe triple.

La singularité conique est telle que l'intersection d'une sphère de dimension quatre, de rayon petit et de centre en un point singulier nous donne deux exemplaires de plan projectif. Les plongements du plan projectif dans cette sphère de dimension quatre sont difféomorphes au plongement qui est défini par cinq fonctions sphériques du second degré sur une 2-sphère (cinq combinaisons linéaires orthonormées x_ix_j, orthogonales à l'unité, définissent dans l'espace des fonctions sur la sphère $x_1^2 + x_2^2 + x_3^2 = 1$ une application paire $S^2 \to S^4$ et par conséquent le plongement $RP^2 \to S^4$).

Il est intéressant de voir comment se comportent les valeurs propres d'une forme quadratique appartenant à une famille générique à deux paramètres à l'approche d'un point singulier en lequel se confondent deux valeurs propres. Un calcul rapide montre que le

graphe de ce couple de valeurs propres est un cône à deux nappes situé au-dessus du plan des paramètres, au voisinage de ce point singulier, et dont le sommet correspond à ce point et chaque nappe à l'une des valeurs propres (fig. 243).

Toute 1-sous-famille générique de la 2-famille considérée a la forme d'une courbe tracée sur le plan des paramètres et ne passant pas

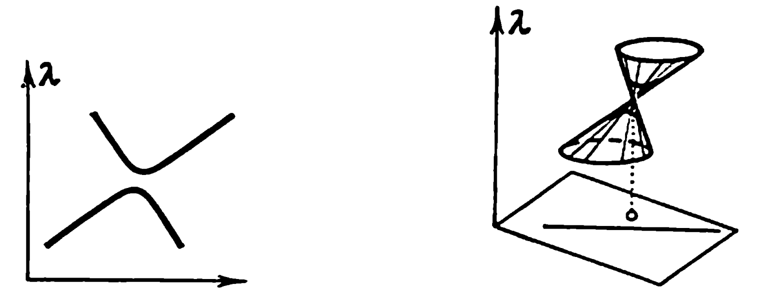


Fig. 243. Fréquences propres de familles à un et deux paramètres de systèmes oscillatoires de forme générale.

par un point singulier. Toute 1-sous-famille contenant un point singulier peut en être débarassée par une petite déformation et la nouvelle 1-famille obtenue passera au voisinage de ce point singulier. Le graphe des valeurs propres situé au-dessus d'une courbe du plan des paramètres passant au voisinage du point singulier est composé des points du cône qui se projettent sur cette courbe. Donc au voisinage du point singulier le graphe est proche d'une hyperbole rappelant un couple de droites concourantes (on obtiendrait un couple de droites si notre famille à un paramètre passait par le point singulier).

Le raisonnement relatif aux valeurs propres des familles à deux paramètres de formes quadratiques explique l'étrange comportement des fréquences propres lorsque l'un des paramètres varie : d'une façon générale (hormis les cas exceptionnels), lorsqu'un paramètre varie, les fréquences propres peuvent se rapprocher l'une de l'autre, sans se dépasser, puis s'éloigner dans des sens opposés.

B. Applications à l'étude des oscillations des milieux continus. Les raisonnements précédents valent également pour l'étude de la dépendance, par rapport aux paramètres, des fréquences propres de divers systèmes mécaniques à nombre fini de degrés de liberté ; cependant, les plus intéressantes sont probablement les applications aux systèmes à une infinité de degrés de liberté qui décrivent les oscillations des milieux continus. Ces applications se basent sur le fait que *les codimensions des variétés des ellipsoïdes à axes multiples sont définies par ces multiplicités et ne dépendent pas de la dimension de l'espace.*

Exemple : l'ensemble des ellipsoïdes de révolution de la variété de tous les ellipsoïdes est de codimension 2 dans un espace de dimen-

sion quelconque ; aussi est-il naturel de considérer que dans la « variété » infinie des ellipsoïdes de l'espace hilbertien de dimension infinie, l'ensemble des ellipsoïdes de révolution est de codimension 2 (et, en particulier, que l'espace des ellipsoïdes ne possédant pas d'axes multiples est connexe).

Certes tout ceci implique une démonstration rigoureuse, mais là n'est pas notre propos. Voyons plutôt à quelles conclusions on aboutit si l'on se hasarde à appliquer les réflexions générales précédentes aux oscillations en milieu continu.

L'énergie cinétique d'un milieu continu occupant un domaine compact D s'exprime en fonction de l'écart u d'un point x par rapport à l'équilibre au moyen de la formule

$$T = \frac{1}{2} \int_D u_t^2 \, dx.$$

Pour fixer les idées nous supposerons que le milieu est une membrane (le domaine D est alors bidimensionnel et l'écart u unidimensionnel). L'énergie cinétique définit une structure euclidienne dans l'espace de configuration du problème (i.e. dans l'espace des fonctions u). L'énergie potentielle est donnée par l'intégrale de Dirichlet

$$U = \frac{1}{2} \int_D (\nabla u)^2 \, dx$$

(au point de vue mathématique ces données figurent dans la définition de la membrane).

Les carrés des fréquences propres de la membrane sont les valeurs propres de la forme quadratique U dans l'espace de configuration sur lequel une métrique est donnée par l'énergie cinétique. Nous *postulerons* qu'à une membrane générique est associée une forme quadratique générique (ceci implique la transversalité de la variété des formes quadratiques associées aux diverses membranes à la variété des formes quadratiques à valeurs propres multiples). Si cette propriété est vérifiée il en découle que

1. Les fréquences propres d'une membrane générique sont distinctes. On peut passer d'une membrane générique à une autre par un chemin continu composé uniquement de membranes à spectre simple.

2. En faisant varier deux paramètres d'une membrane, on peut amener deux valeurs propres à coïncider ; pour obtenir une fréquence triple il faut disposer de 5 paramètres indépendants, une quadruple — de neuf paramètres, etc.

3. Si l'on part d'une membrane à spectre simple et qu'on la déforme continûment pour passer à une autre membrane à spectre simple suivant un chemin général quelconque, alors au terme de ce prolongement la k-ème (par ordre de grandeur) fréquence propre de la membrane initiale se transforme toujours, indépendamment du

chemin suivi par la déformation, en la k-ème (par ordre de grandeur) fréquence propre de la membrane finale; le prolongement des fonctions propres dépend en général du chemin suivi par la déformation (plus exactement, la modification du chemin peut changer le signe de la fonction propre obtenue).

En particulier, si partant d'une membrane à spectre simple et la déformant on décrit un chemin fermé dans l'espace des membranes, i.e. l'on revient à la membrane initiale après avoir parcouru un ensemble de membranes à spectre multiple (qui, on le sait, est de codimension 2), alors la k-ème fréquence propre prendra sa valeur initiale, tandis que la k-ème fonction propre est susceptible de changer de signe.

C. Effet des symétries sur la multiplicité du spectre. Le spectre multiple est exceptionnel dans les systèmes généraux, mais il se manifeste et il est alors impossible à éliminer par une petite déformation si le système donné est symétrique et les déformations conservent cette symétrie.

Soit par exemple dans le plan un système de trois masses égales mobiles placées aux sommets d'un triangle équilatéral et liées entre elles et au centre du triangle par des ressorts équivalents. Ce système possède un axe de répétition d'ordre trois. Donc sur son espace de configuration (qui est de dimension 6) évolue un opérateur linéaire g dont le cube est égal à 1 et qui respecte aussi bien la structure euclidienne (définie par l'énergie cinétique) que l'ellipsoïde de l'espace de configuration qui définit l'énergie potentielle.

Il en résulte que *cet ellipsoïde doit être de révolution.* En effet, si g est l'opérateur de symétrie dans l'espace de configuration et ξ un vecteur porté par un axe principal de l'ellipsoïde, alors l'axe $g\xi$ est également principal (car la rotation g applique l'ellipsoïde sur lui-même).

Deux cas sont possibles pour le vecteur $g\xi$: soit que $g\xi = \xi$, soit que $g\xi$ et ξ sont linéairement indépendants et font un angle de 120° dans l'espace de configuration. Dans le deuxième cas, le plan engendré par ξ et $g\xi$ est composé uniquement d'axes principaux et par conséquent la valeur propre correspondant à un tel axe est au minimum double.

Ce raisonnement montre que les oscillations propres d'un système possédant un axe de répétition d'ordre trois peuvent être de deux types: celles qui se conservent par une rotation de 120° ($g\xi = \xi$) et celles qui, par cette rotation, se transforment en une oscillation propre indépendante de même fréquence ($g\xi$ et ξ sont indépendants). Dans le deuxième cas il s'introduit même trois formes d'oscillations propres de fréquence commune (ξ, $g\xi$ et $g^2\xi$) dont deux seulement sont indépendantes:

$$\xi + g\xi + g^2\xi = 0,$$

puisque la somme de trois vecteurs coplanaires de même module faisant entre eux des angles de 120° est nulle.

Au total notre système possède 6 oscillations propres. Procédons comme suit pour définir lesquelles d'entre elles sont du premier type (symétrique) et lesquelles du second (non symétrique). Considérons le cas limite où chaque masse oscille indépendamment des autres.

Nous pouvons alors choisir dans l'espace de configuration une base orthonormée composée de six oscillations propres à raison de deux par point, telle qu'un point se déplace et les deux autres non. Désignons par ξ_i et η_i les vecteurs propres de fréquences propres a et b respectivement correspondant au i-ème point, et soient x_i, y_i leurs coordonnées dans la base orthonormée ξ_i, η_i. L'énergie potentielle s'écrit alors

$$U=\frac{1}{2}(a^2x_1^2+b^2y_1^2)+\frac{1}{2}(a^2x_2^2+b^2y_2^2)+\frac{1}{2}(a^2x_3^2+b^2y_3^2).$$

L'opérateur de symétrie g permute les axes de coordonnées comme suit :

$$g\xi_1=\xi_2,\quad g\xi_2=\xi_3,\quad g\xi_3=\xi_1,$$
$$g\eta_1=\eta_2,\quad g\eta_2=\eta_3,\quad g\eta_3=\eta_1.$$

Nous sommes à présent en mesure de représenter notre espace de dimension six sous forme d'une somme orthogonale directe de deux droites et de deux 2-plans, invariants par l'opérateur de symétrie g. Plus exactement, les droites invariantes sont définies par les vecteurs directeurs

$$\xi_1+\xi_2+\xi_3,\quad \eta_1+\eta_2+\eta_3,$$

quant aux 2-plans invariants, ce sont leurs orthocompléments relativement aux espaces engendrés respectivement par les vecteurs de base unitaires ξ_i et η_i.

De plus, la première droite est la direction de l'oscillation propre symétrique de fréquence a et la seconde, de fréquence b. De la même façon tout vecteur du premier plan est direction de l'oscillation propre de fréquence a qui, par une rotation de 120°, se transforme en une oscillation indépendante de même fréquence; pour tous les vecteurs du second plan l'oscillation n'est pas symétrique non plus et est de fréquence b.

Donc dans le cas dégénéré des trois points indépendants *il existe deux oscillations propres indépendantes de type symétrique et quatre de type non symétrique, les dernières étant classées en deux couples. Dans chaque couple les oscillations possèdent une fréquence propre commune et se déduisent l'une de l'autre par une rotation du plan des trois points de* 120°.

Nous affirmons maintenant que ce qui précède est valable pour toute loi d'interaction entre nos points pourvu que cette interaction

soit symétrique, i.e. que l'énergie potentielle du système se conserve par une rotation du plan de 120°.

En effet, prenons un système quelconque ayant six oscillations propres $\zeta_1, \ldots, \zeta_6$, deux à deux orthogonales. Chaque vecteur ζ_i de l'espace de configuration soit reste fixe soit tourne de 120° par l'opérateur g. Or les vecteurs fixes par l'opérateur g forment un 2-plan et ceux qui tournent de 120°, un espace de dimension quatre. Donc à deux des six vecteurs correspondent des oscillations symétriques; quant aux quatre autres, ils sont contenus dans l'espace de dimension quatre des vecteurs tournant de 120°, perpendiculaire aux deux premiers. Considérons l'un de ces quatre vecteurs, appliquons-lui l'opérateur g et convenons de dire du vecteur obtenu qu'il est couplé avec la direction initiale de l'oscillation propre. Puis dans l'orthocomplément du plan obtenu relativement à l'espace de dimension quatre, prenons un vecteur quelconque et couplons-le avec son image par l'opérateur g. Nous obtenons ainsi un système composé de six oscillations propres jouissant des propriétés annoncées.

Donc *dans un système générique de trois points du plan, possédant un axe de répétition d'ordre trois, il existe quatre fréquences propres distinctes dont deux simples et deux doubles. Ceci étant, à chaque fréquence propre simple est associée une oscillation propre symétrique, et à chaque fréquence double, trois oscillations propres déduites l'une de l'autre par une rotation de* 120° *et de somme nulle* (*de sorte que seules deux d'entre elles sont indépendantes*).

Exercice. Classer les oscillations propres d'un système possédant les mêmes éléments de symétrie qu'un triangle équilatéral (admettant des rotations de 120° et des symétries par rapport aux médiatrices).

Exercice. Classer les oscillations propres d'un système dont le groupe des symétries est le groupe des 24 rotations du cube.

Réponse. Les oscillations sont de cinq types. Chaque oscillation peut donner par les rotations un système de 8, ou 6, ou 4, ou 2, ou 1 oscillation indépendante (dans le dernier cas l'oscillation est totalement symétrique).

Remarque. Un appareil spécial (appelé théorie des représentations des groupes) a été conçu pour la classification des oscillations dans les systèmes à groupes de symétries quelconques. Voir G. Lubarsky, *Théorie des groupes et son application à la physique*, (en russe), Moscou, 1958, dans lequel figurent les tables nécessaires.

D. Comportement des fréquences d'un système symétrique par une variation des paramètres respectant la symétrie. Supposons que notre système symétrique dépend d'une façon générale d'un certain nombre de paramètres dont la variation respecte la symétrie. Alors les fréquences propres de multiplicités diverses dépendront également des paramètres et nous assisterons à des collisions entre les fréquences propres. Nous nous limiterons à la formulation du résultat pour le cas simple d'un système possédant un axe de répétition d'ordre trois (la réponse est la même si l'axe de répétition est d'ordre

$n \geqslant 3$). On trouvera les détails dans les articles: V. Arnold, *Modes et quasi-modes* (en russe). « Analyse fonctionnelle et ses applications », 6 : 2 (1972), 12-20; V. Karpouchkine, *Sur l'asymptotique des valeurs propres des variétés symétriques et sur les « plus probables » représentations des groupes finis* (en russe). « Courrier de l'Université de Moscou », série math., n° 2, 1974, 9-13.

Les oscillations propres d'un système quelconque possédant un axe de répétition d'ordre trois sont de deux types : celles qui sont symétriques et celles qui se transforment en oscillations indépendantes de l'initiale par une rotation de 120°. Si un système générique possède un axe de répétition d'ordre trois (un tel système ne possède pas d'autres symétries) toutes les fréquences propres du premier type sont simples et celles du second, doubles.

Si par ailleurs un système est symétrique pour toutes les valeurs du seul paramètre dont il dépend d'une façon générale, *alors lorsque celui-ci varie les fréquences propres des oscillations symétriques ne se heurtent pas, et les fréquences doubles des oscillations non symétriques ne se décomposent pas. De plus, les fréquences propres doubles des oscillations non symétriques ne se heurtent pas non plus. Cependant les fréquences propres des oscillations symétriques et non symétriques varient indépendamment les unes des autres de sorte que pour certaines valeurs du paramètre une fréquence propre d'une oscillation symétrique peut heurter (et passer à travers) une fréquence propre (double) d'une oscillation non symétrique.*

Pour que deux fréquences propres des oscillations symétriques se heurtent il faut faire varier au minimum deux paramètres ; pour les oscillations non symétriques, au moins trois paramètres.

En général, pour que dans une famille générique de systèmes possédant des axes de répétition d'ordre trois on trouve des systèmes correspondant à la collision entre i fréquences propres simples (i oscillations symétriques) et j doubles (j oscillations non symétriques), le nombre des paramètres de la famille doit être au minimum égal à

$$(i-1)(i+2)/2+j^2.$$

Appliquons ce que nous venons de dire aux oscillations de membranes symétriques. Nous supposerons qu'à une membrane générique invariante par une rotation de 120° est associé un ellipsoïde générique appartenant à l'espace des ellipsoïdes de l'espace de configuration, invariants par une transformation de l'espace de configuration induite par la rotation de la membrane.

L'énoncé exact de cette hypothèse est : l'application de l'espace des membranes symétriques dans l'espace des ellipsoïdes symétriques est transversale aux variétés des ellipsoïdes ayant un nombre distinct d'axes multiples pour toutes les membranes à l'exception d'un ensemble de codimension infinie.

Si cette hypothèse est vraie l'on déduit à propos des oscillations des membranes symétriques :

1. Pour les membranes génériques invariantes par une rotation de 120° asymptotiquement le tiers des fréquences propres (comptées avec leurs multiplicités) sont simples et les oscillations propres correspondantes sont invariantes par une rotation de 120°. Les autres fréquences propres sont doubles et à chacune d'elles sont associées trois fonctions propres, de somme nulle, se déduisant l'une de l'autre par une rotation de 120°.

2. Dans des familles génériques à un paramètre de telles membranes symétriques peuvent coïncider, pour des valeurs distinctes des paramètres, une fréquence simple avec une fréquence double, mais jamais deux fréquences simples ni deux fréquences doubles.

3. Le nombre minimum de paramètres d'une famille de membranes où l'on obtient des collisions plus complexes (qu'on ne peut éviter par un petit déplacement respectant la symétrie) entre des fréquences propres est donné par la formule

$$\sum_{i,j}[(i-1)(j+2)/2+j^2]\,\nu_{ij},$$

où ν_{ij} est le nombre de points de collisions entre i fréquences simples et j doubles.

En particulier, si une membrane circulaire est soumise à l'effet d'une petite déformation générique respectant la rotation de 120°, le tiers des valeurs propres doubles (correspondant aux fonctions propres à partie azimutale $\cos 3k\varphi$ et $\sin 3k\varphi$) se découplent aussitôt. Si l'on poursuit la déformation à un paramètre, les fréquences propres simples et doubles peuvent passer l'une à travers l'autre, mais il n'y aura collision ni entre deux simples ni entre deux doubles fréquences.

E. Discussion. Les considérations génériques et de symétrie sont importantes en ce sens qu'elles permettent de se faire une idée sur les problèmes dont la solution exacte est difficile à obtenir.

En particulier, pour aucune membrane ou presque on ne connaît l'allure des oscillations propres. Ce qui n'empêche qu'à partir de raisonnements généraux on puisse dire quelque chose sur les multiplicités des valeurs propres par exemple.

L'étude des oscillations de haute fréquence des milieux continus est primordiale dans de nombreux domaines (optique, acoustique, etc.). Des méthodes spéciales ont été élaborées pour déterminer approximativement l'allure des oscillations propres. Une méthode (appelée *asymptotique quasi classique*) consiste à chercher l'oscillation sous forme d'une onde de petite longueur localement proche d'un harmonique simple, dont l'amplitude et la direction du front varient légèrement d'un point à l'autre.

L'analyse (que nous omettrons) montre que dans certains cas on peut construire des solutions approchées pour l'équation des fonctions propres jouissant des propriétés indiquées. Ces solutions sont approchées en ce sens qu'elles vérifient presque l'équation pour les fonctions propres (et non qu'elles sont proches des fonctions propres).

En particulier, si la membrane possède la forme d'un triangle équilatéral dont on a fortement arrondi les sommets, on peut alors construire une solution approchée qui ne différera sensiblement de la solution nulle qu'au voisinage d'un sommet du triangle. Les physiciens appellent cette solution approchée analogue ondulatoire d'un rayon en mouvement sur une hauteur du triangle ; ce rayon est une trajectoire stable *) d'un billard dont la forme est celle de la membrane (voir Appendice sur les asymptotiques à ondes courtes).

Pour des raisons de symétrie et du cas générique il résulte que les membranes génériques ayant un axe de répétition d'ordre trois ne possèdent pas de véritables oscillations propres de l'espèce décrite. Supposons en effet qu'une des oscillations propres de la membrane est concentrée à proximité d'une médiatrice (mais pas au centre de la membrane). Alors en la faisant pivoter de 120° puis de 240° on obtient trois oscillations propres de même fréquence propre. Ces trois oscillations sont indépendantes (ceci découle de la non-nullité de leur somme). Donc la fréquence propre est triple, or de tels systèmes génériques possédant un axe de répétition d'ordre trois n'existent pas.

On remarque donc qu'il est vain d'essayer de construire une asymptotique de haute fréquence rigoureuse pour les fonctions propres : le mieux que l'on puisse espérer c'est obtenir des formules approchées pour des oscillations presque propres. Une telle oscillation presque propre risque de s'écarter fortement d'une véritable oscillation propre, cependant si l'on se donne une condition initiale qui lui corresponde, alors le mouvement rappellera une onde stationnaire (oscillation propre) sur un long intervalle de temps.

Un exemple d'oscillation presque propre est le mouvement d'un pendule couplé à un autre identique par un ressort lâche. Si à la date initiale nous mettons un pendule en mouvement, l'autre étant fixe, on constate que le premier oscillera seul pendant longtemps et ses oscillations seront presque propres. Dans le cas de véritables oscillations propres les amplitudes des deux pendules sont égales.

La géométrie de la membrane et les propriétés de ses oscillations propres ont fait l'objet d'intenses recherches par nombre d'auteurs et notamment par H. Weyl, S. Minakshisundaram et A. Pleijel, A. Selberg, J. Milnor, M. Kac,

---

*) La condition de stabilité linéaire d'une trajectoire de billard s'écrit

$$(r_1 + r_2 - l)(r_1 - l)(r_2 - l) > 0,$$

où l est la longueur d'une portion de trajectoire, r_1 et r_2 les rayons de courbure de la paroi aux extrémités.

I. Singer, N. Mckean, M. Berger, Y. Colin de Verdière, J. Chazarain, J. J. Duistermaat, V. Lasoutkine, A. Shnirelman, S. Moltchanov.

A la question simple « peut-on déterminer la forme d'un tambour à l'ouïe? » la réponse est négative : il existe des variétés riemanniennes non isométriques à spectre commun. Par ailleurs, certaines propriétés de la variété sont restituées d'après le spectre des valeurs propres de l'opérateur de Laplace et d'après les propriétés des fonctions propres (il est possible de restituer par exemple la collection complète des longueurs des géodésiques fermées).

APPENDICE 11

ASYMPTOTIQUES À ONDES COURTES

La description de la propagation de la lumière en optique géométrique à l'aide des rayons (i.e. à l'aide des équations canoniques de Hamilton) ou des fronts d'ondes (i.e. à l'aide des équations de Hamilton-Jacobi) est approchée au point de vue de l'optique physique. En optique physique la lumière est une onde électromagnétique, et l'optique géométrique une première approximation qui décrit bien le phénomène seulement lorsque la longueur d'onde est petite devant les dimensions des corps considérés.

L'équivalent mathématique de ces notions physiques sont les formules asymptotiques pour les solutions des équations différentielles correspondantes, formules qui donnent une approximation d'autant meilleure que la fréquence des oscillations est élevée (i.e. que les ondes sont plus courtes). Ces formules asymptotiques s'écrivent en termes de rayons (i.e. de mouvements dans un certain système dynamique hamiltonien) ou de fronts (i.e. de solutions des équations de Hamilton-Jacobi).

Une telle asymptotique à ondes courtes existe pour les solutions de nombreuses équations de physique mathématique décrivant tous les processus ondulatoires possibles. On lui attribue des noms différents selon le domaine de la physique ou des mathématiques considéré. En mécanique quantique, par exemple, on l'appelle approximation quasi classique et on la recherche par la méthode WKBJ (G. Wentzel, H. A. Kramers, L. Brillouin, H. Jeffreys), bien que cette approximation ait déjà été utilisée par Liouville, Green, Stokes et Rayleigh.

La construction de l'asymptotique à ondes courtes est basée sur l'observation locale d'une série d'ondes presque sinusoïdales, dont, cependant, l'amplitude et la direction des fronts varient lentement d'un point à l'autre. En portant formellement une fonction de cette forme dans l'équation aux dérivées partielles décrivant le processus ondulatoire on obtient (en première approximation pour une petite longueur d'onde) l'équation de Hamilton-Jacobi pour les fronts d'ondes. Les approximations suivantes permettent de

définir également la dépendance de l'amplitude des oscillations par rapport au point.

Certes il faut justifier mathématiquement cette procédure. La formulation exacte et la démonstration des théorèmes respectifs ne sont pas de tout repos. Les plus grosses difficultés proviennent des caustiques (autrement: les points focaux, ou les points conjugués, ou les points de retour).

Les caustiques sont les enveloppes de familles de rayons; on peut les observer sur un mur éclairé par des rayons réfléchis par une surface courbe lisse. Si les rayons intervenant dans la description des ondes se coupent et forment des caustiques, alors au voisinage de ces dernières les formules de l'asymptotique à ondes courtes doivent être légèrement modifiées. Plus exactement, la phase des oscillations le long de chaque rayon subit une discontinuité ordinaire (d'un quart d'onde) à chaque passage du rayon près de la caustique.

Tous ces phénomènes sont commodes à décrire dans le langage de la géométrie des sous-variétés lagrangiennes de l'espace des phases correspondant et de leurs projections sur l'espace de configuration.

Les caustiques s'interprètent alors comme les singularités de la projection de la variété lagrangienne, qui définit la famille de rayons, de l'espace des phases sur l'espace de configuration. Donc les formes normales des singularités des projections lagrangiennes qu'on trouvera dans l'Appendice 12 fournissent en particulier la classification des singularités des caustiques engendrées par les systèmes de rayons génériques.

Dans le présent Appendice sont proposées (sans démonstration) les formules simples de l'asymptotique à ondes courtes pour l'équation de Schrödinger de la mécanique quantique. On trouvera un exposé plus détaillé dans les ouvrages suivants:

1. J. Heading — An introduction to phase integral methods. London, 1962.

2. V. Maslov — Théorie des perturbations et méthodes asymptotiques (en russe). Université de Moscou, 1965.

3. V. Arnold — Sur une classe caractéristique figurant dans les conditions de quantification (en russe). « Analyse fonctionnelle et ses applications », t. 1, fasc. 1 (1967), 1-14.

4. L. Hörmander — Fourier integral operators. « Acta Mathematica » 127, 1-2 (1971), 119.

A. Approximation quasi classique pour les solutions de l'équation de Schrödinger. On appelle *équation de Schrödinger* pour une particule en mouvement dans un champ de potentiel U de l'espace euclidien l'équation en la fonction complexe $\psi(q, t)$

$$ih\frac{\partial\psi}{\partial t} = -\frac{h^2}{2}\Delta\psi + U(q)\psi, \quad q \in \mathbf{R}^n, \quad t \in \mathbf{R}.$$

Ici h est une constante réelle qui figure le petit paramètre dans le problème considéré, Δ l'opérateur de Laplace.

Supposons que la condition initiale est de la forme

$$\psi|_{t=0} = \varphi(q)\, e^{\frac{i}{h} s(q)},$$

où la fonction différentiable φ est non nulle seulement à l'intérieur d'un domaine borné. Nous donnerons plus bas la formule asymptotique (pour $h \to 0$) pour la solution de l'équation de Schrödinger qui réalise cette condition initiale.

Considérons tout d'abord le mouvement d'une particule classique dans un champ de potentiel U, i.e. soit l'équation de Hamilton

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q} \quad \text{où } H = \frac{1}{2} p^2 + U(q)$$

dans un espace des phases de dimension $2n$. Les solutions de ces équations définissent un flot (moyennant certaines conditions sur le potentiel que nous supposons réalisées; ces conditions interdisent à la solution de s'éloigner à l'infini en un temps fini).

Associons à notre condition initiale une sous-variété lagrangienne dans l'espace des phases (i.e. une variété ayant la même dimension que l'espace de configuration et sur laquelle est identiquement nulle la 2-forme $dp \wedge dq$ qui définit la structure symplectique sur l'espace des phases). Plus exactement, nous définissons l'« impulsion » correspondant à notre condition initiale comme le gradient de phase, i.e. posons

$$p(q) = \partial s / \partial q.$$

Lemme. *Quelle que soit la fonction différentiable s, le graphe de la fonction $p(q)$ dans l'espace des phases $\mathbf{R}^{2n} = \{(p, q)\}$ est une variété lagrangienne. Réciproquement, si une variété lagrangienne se projette de façon unique sur un q-espace (i.e. est un graphe), alors elle est définie par une fonction génératrice s au moyen de la formule précédente.*

Désignons par M la variété lagrangienne construite à l'aide de la condition initiale (à l'aide de la fonction s). Le flot g^t transforme dans le temps t la variété M en une autre variété $g^t M$. Cette nouvelle variété est également lagrangienne. Cependant cela n'est forcément pas le cas pour les grands t (fig. 244).

En d'autres termes, plusieurs points de $g^t M$ peuvent se projeter en un même point Q de l'espace de configuration. Supposons que ces points sont en nombre fini et qu'ils sont tous non dégénérés (i.e. n'est pas dégénérée la dérivée de l'application projective de $g^t M$ sur l'espace de configuration en chacun des points qui se projettent en Q).

La condition de non-dégénérescence est réalisée pour presque tous les points Q sauf pour un ensemble de points singuliers de mesure nulle dans l'espace de configuration. Dans le cas général cet ensemble est une surface dont la dimension

est d'une unité inférieure à celle de l'espace de configuration. Cette surface qui joue le rôle de caustique dans notre problème est susceptible elle-même de présenter des singularités assez compliquées.

Les points de $g^t M$ qui se projettent sur le point donné Q sont les images, par les transformations du flot, de certains points de la variété lagrangienne initiale M. En d'autres termes, au point Q aboutissent dans le temps t plusieurs trajectoires de la particule classique, trajectoires dont les conditions initiales appartiennent à M.

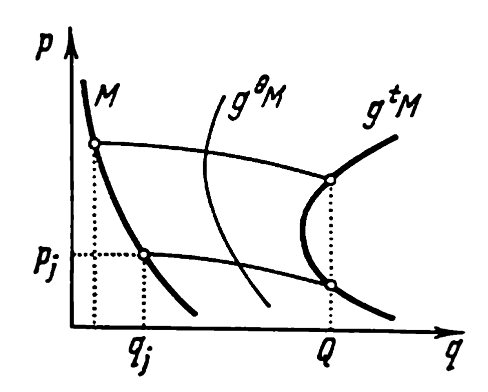


Fig. 244. Transformation des variétés lagrangiennes par le flot.

Désignons ces conditions initiales par (p_j, q_j) et l'action du flot le long de la trajectoire issue du point (p_j, q_j) par S_j. Plus exactement, posons

$$S_j(Q, t) = s(q_j) + \int_0^t L\,d\theta, \quad \text{où} \quad L = \frac{\dot{q}^2}{2} - U(q) \text{ et } g^\theta(p_j, q_j) = $$
$$= (p(\theta), q(\theta)).$$

Alors lorsque $h \to 0$ la solution de l'équation de Schrödinger dont les fonctions s et φ sont données par une condition initiale oscillante admet pour asymptotique

$$\psi(Q, t) = \sum_j \varphi(q_j) \left| \frac{DQ}{Dq_j} \right|^{-1/2} e^{\frac{i}{h} S_j(Q, t) - \frac{i\pi}{2}\mu_j} + O(h),$$

où μ_j est un entier (indice de Morse) *qui sera défini plus bas.*

Pour bien saisir le sens de cette formule traitons d'abord le cas où l'intervalle de temps t est petit. La somme se réduit alors à un seul terme puisque la variété lagrangienne $g^t M$ déduite de M par le flot pendant un temps t petit se projette de façon unique sur l'espace de configuration. En d'autres termes, de toute la famille de particules correspondant à la condition initiale de l'équation de Schrödinger, une seule arrive en Q au bout d'un temps t petit.

Pour les t petits l'indice de Morse est nul (comme il résultera de sa définition). Donc la fonction $\psi(Q, t)$ comme la condition initiale oscille rapidement.

Ceci étant, la fonction S qui définit les fronts d'ondes à l'instant t n'est autre chose que la valeur à l'instant t de la solution de l'équation de Hamilton-Jacobi dont la condition initiale est définie par une fonction s déterminant les fronts d'ondes à l'instant initial. L'amplitude des ondes à l'instant t au point Q se déduit de leur amplitude au point de départ de la trajectoire aboutissant en Q par multiplication par un certain facteur. Ce facteur est choisi de manière que lorsque les particules sont en mouvement l'intégrale du carré du module de la fonction ψ étendue au domaine de l'espace de configuration, adhérence des particules correspondant à notre condition initiale, soit fixe. (On suppose qu'à l'instant initial a été choisi un domaine quelconque de l'espace de configuration, puis on considère sur la variété M les points représentatifs dont les projections sur l'espace de configuration sont situées dans le domaine choisi, ensuite leurs images par le flot et enfin les projections de ces images sur l'espace de configuration, projections qui forment un domaine, « rempli de particules à l'instant t ».)

B. Indices de Morse et de Maslov. Le nombre μ_j est par définition le nombre de points focaux à la variété M, compris dans le segment $[0, t]$ de la courbe de phase issue du point (p_j, q_j).

On définit un point focal à M de la manière suivante. Nous avons choisi le point Q tel que la condition de non-dégénérescence soit réalisée par l'application projective de la variété lagrangienne g^tM. Cependant si nous considérons la courbe de phase issue du point (p_j, q_j) toute entière, alors à certains instants θ compris entre 0 et t la condition de non-dégénérescence est susceptible de ne pas se réaliser au point $(p(\theta), q(\theta))$ de la variété lagrangienne $g^\theta M$. Ces points sont appelés *points focaux* à la variété M le long de la courbe de phase considérée.

Remarquons que les définitions d'un point focal à M et de l'indice de Morse ne dépendent pas de l'équation de Schrödinger, elles relèvent simplement de la géométrie du flot dans le fibré cotangent à l'espace de configuration (ou, ce qui revient au même, du calcul variationnel). En particulier, pour variété lagrangienne M on peut prendre une fibre du fibré cotangent passant par le point (p_0, q_0) (cette fibre est définie par la condition $q = q_0$).

Dans ce cas le point focal à M situé sur la courbe de phase issue de (p_0, q_0) est par définition *conjugué* du point initial (plus exactement, la projection de ce point focal sur l'espace de configuration est appelée point conjugué du point q_0 le long de l'extrémale issue du point q_0 avec l'impulsion p_0 dans cet espace). Dans le cas encore plus particulier du mouvement sur une géodésique d'une variété riemannienne, on appelle point focal à une fibre du fibré cotangent le point conjugué du point initial d'une géodésique le long d'elle. Exemple : le pôle sud de la sphère est conjugué du pôle nord le long d'un méridien quelconque.

L'indice de Morse d'un segment de géodésique, indice qui est égal au nombre de points conjugués de l'origine, joue un rôle important en calcul variationnel. Plus exactement, considérons la différentielle seconde de l'action comme une forme quadratique sur l'espace des variations de la géodésique étudiée (à extré-

mités fixes). Alors l'indice négatif d'inertie de cette forme quadratique est égal à l'indice de Morse (voir par exemple J. Milnor, *Morse Theory*, Princeton, 1963).

Donc la géodésique réalise le minimum de l'action avant le premier point conjugué, justifiant ainsi la dénomination « principe de moindre action » de certains principes variationnels de mécanique.

Signalons que lorsqu'on calcule l'indice de Morse, il faut tenir compte des multiplicités des points focaux (la multiplicité d'un point focal générique est 1).

L'indice de Morse est un cas particulier de l'*indice de Maslov* qui se définit indépendamment du flot pour des courbes quelconques sur une variété lagrangienne du fibré cotangent au-dessus de l'espace de configuration.

Considérons l'application projective de la variété lagrangienne M sur l'espace de configuration de dimension n. C'est une application différentiable de variétés de même dimension. Elle peut présenter des points singuliers, i.e. des points en lesquels le rang de sa dérivée baisse et dans le voisinage desquels elle n'est pas un difféomorphisme.

On constate que *d'une façon générale l'ensemble des points singuliers est de dimension* $(n-1)$ *et est constitué de l'union de la variété différentiable de dimension* $(n-1)$ *des points singuliers simples en lesquels le rang baisse de* 1 *et d'une collection finie de variétés de dimension* $\leqslant n-3$.

Ici le « d'une façon générale » signifie que l'on peut réaliser les propriétés indiquées par une déformation aussi petite que l'on veut de M conservant la structure lagrangienne de M.

Il est essentiel de signaler que parmi les sous-ensembles de rangs divers en lesquels on partage l'ensemble des points singuliers, il n'en existe aucun qui soit de dimension $(n-2)$. Aux points singuliers simples qui forment une variété de dimension $(n-1)$ succèdent les points où le rang baisse de deux unités et ces points constituent une variété de dimension $(n-3)$. La projection de l'ensemble des points singuliers sur l'espace de configuration (la caustique) est composée de sous-ensembles de toutes les dimensions de 0 à $(n-1)$ sans exception.

On constate d'autre part que la *variété de dimension* $(n-1)$ *des points singuliers simples a deux côtés sur la variété de Lagrange*, plus exactement on peut convenir de l'orientation des normales en tous ses points en procédant comme suit.

Soit un point singulier simple quelconque d'une variété lagrangienne.

Considérons un système de coordonnées $q_1, \ldots, q_n$ au voisinage de la projection de ce point sur l'espace de configuration. Soient $p_1, \ldots, p_n$ les coordonnées correspondantes dans les fibres du fibré cotangent. Au voisinage du point singulier la variété peut être considérée comme le graphe d'un vecteur-fonction $(q_1, p_2, \ldots, p_n)$ des variables $(p_1, q_2, \ldots, q_n)$ (ou d'un vecteur-fonction de la

même forme dans lequel la coordonnée privilégiée sera non pas la première mais une autre quelconque).

Les points singuliers proches du point envisagé se déterminent alors à partir de la condition $\partial q_1/\partial p_1 = 0$. Pour les variétés lagrangiennes génériques cette dérivée change de signe au passage d'un côté de la variété des points singuliers sur l'autre au voisinage considéré du point singulier simple. Le *côté positif* sera celui où cette dérivée est positive.

A remarquer qu'il faut démontrer que cette définition du sens positif concorde au voisinage des divers points. Bien plus, il faut démontrer que le sens positif est intrinsèquement défini au voisinage d'un point, i.e. ne dépend pas du système de coordonnées. Ces démonstrations se font par des calculs directs (voir l'article cité plus haut dans l'« Analyse fonctionnelle »).

L'indice de Maslov d'une courbe orientée d'une variété lagrangienne se définit maintenant comme le nombre de passages du côté négatif au côté positif de la variété des singularités diminué du nombre de passages inverses. On suppose que les extrémités de la courbe ne sont pas singulières et que cette courbe ne coupe que la variété des points singuliers simple et ce sous des angles non nuls. Définissons maintenant l'indice d'une courbe quelconque dont les extrémités sont des points non singuliers: il suffit pour cela de l'approcher par une courbe qui ne coupe que la variété des points singuliers simples et ce sous des angles non nuls. On montre que l'indice ne dépend pas du choix de la courbe approximisante.

Exercice. Trouver l'indice du cercle $p = \cos t$, $q = \sin t$ orienté par le paramètre t, $0 \leqslant t \leqslant 2\pi$, sur la variété lagrangienne $p^2 + q^2 = 1$ du plan de phase.

Réponse. $+2$.

Enfin l'indice de Morse d'une courbe de phase de $\mathbf{R}^{2n}$ peut maintenant être défini comme l'indice de Maslov d'une courbe d'une variété lagrangienne de dimension $(n + 1)$ d'un espace des phases de dimension $(2n + 2)$ dûment choisi. Les coordonnées de cet espace sont $(p_0, p; q_0, q)$, où $(p, q) \in \mathbf{R}^{2n}$. Si l'on pose ici $q_0 = t$, $p_0 = -H(p, q)$ et que l'on fasse parcourir au point (p, q) la variété lagrangienne de dimension n de $\mathbf{R}^n$ déduite de la variété initiale par le flot pendant le temps t, alors lorsque t variera, les points obtenus dans $\mathbf{R}^{2n+2}$ balayeront une variété lagrangienne de dimension $(n + 1)$. Le graphe du mouvement du point représentatif par le flot peut être regardé comme une courbe sur cette variété lagrangienne de dimension $(n + 1)$. Il est aisé de vérifier que l'indice de Maslov de ce graphe est l'indice de Morse de la courbe de phase initiale.

C. Indices des courbes fermées. Les indices des courbes fermées tracées sur des sous-variétés lagrangiennes d'un espace des phases vectoriel peuvent être également calculés à l'aide d'une structure

complexe. Munissons l'espace des phases vectoriel $\mathbf{R}^{2n} = \{(p, q)\}$ en plus de la structure symplectique $dp \wedge dq$, d'une structure euclidienne (de carré scalaire $p^2 + q^2$) et d'une structure complexe définie par le produit par l'unité imaginaire

$$I: \mathbf{R}^{2n} \to \mathbf{R}^{2n},\ I(p, q) = (-q, p);\ z = p + iq,\ \mathbf{C}^n = \{z\}.$$

Ces trois structures sont liées par la relation

$$[x, y] = (Ix, y),$$

où le crochet désigne le produit scalaire gauche.

Les transformations linéaires de l'espace des phases qui conservent deux quelconques de ces structures (et partant toutes les trois) s'appellent *unitaires*. Ces transformations envoient les plans lagrangiens dans des plans lagrangiens.

Tout plan lagrangien peut se déduire d'un plan quelconque (par exemple d'un plan réel de $\mathbf{R}^n$ défini par l'équation $q = 0$) par une transformation unitaire. Ceci étant, deux transformations unitaires A, B quelconques qui envoient un plan réel dans un même plan lagrangien diffèrent l'une de l'autre par une transformation unitaire qui est orthogonale et réelle:

$$B = AC, \text{ où } C\mathbf{R}^n = \mathbf{R}^n.$$

Inversement, toute transformation orthogonale préliminaire ne change pas l'image d'un plan réel par une transformation unitaire.

Remarquons maintenant que le déterminant d'une transformation orthogonale est égal à ± 1. Donc le *carré du déterminant* d'une transformation unitaire qui envoie un plan réel dans un plan lagrangien donné ne dépend que du plan lagrangien et en aucune façon du choix de la transformation unitaire.

Ces remarques préliminaires faites, revenons à notre variété lagrangienne et à la courbe fermée orientée traçée sur elle. En chacun de ses points cette courbe admet un plan tangent à la variété lagrangienne dans l'espace vectoriel symplectique. Le carré du déterminant de la transformation unitaire qui envoie un plan réel dans un plan tangent est un nombre complexe de module égal à l'unité. Lorsque le point représentatif se déplace sur la courbe fermée, ce nombre complexe varie. Quand le point représentatif effectue un tour complet sur la courbe fermée, le carré du déterminant contourne un nombre entier de fois l'origine des coordonnées du plan de la variable complexe, qui est orienté de 1 à i. Ce nombre entier représente précisément l'indice de la courbe fermée considérée.

Les indices des courbes fermées interviennent dans les formules asymptotiques des problèmes stationnaires (des oscillations propres). Supposons que le flot correspondant au potentiel U admet une variété lagrangienne invariante située sur le niveau d'énergie $H = E$.

Alors l'équation

$$\frac{1}{2}\Delta\psi = \lambda^2 (U(q) - E)\,\psi$$

admet une série de valeurs propres $\lambda_N \to \infty$ d'asymptotique $\lambda_N = \mu_N + O(\mu_N^{-1})$ si pour tous les contours γ fermés tracés sur notre variété lagrangienne est réalisée l'identité

$$\frac{2\mu_N}{\pi} \oint_\gamma p\,dq \equiv \operatorname{ind}\gamma \pmod 4.$$

Dans le cas unidimensionnel la variété lagrangienne est un cercle dont l'indice est égal à 2 et la précédente formule se transforme en la « condition de quantification »:

$$\mu_N \oint_\gamma p\,dq = 2\pi\left(N + \frac{1}{2}\right).$$

Les fonctions propres qui correspondent à ces valeurs propres sont également liées à la variété lagrangienne, mais ce lien est compliqué. En effet, on ne peut pas établir de formules asymptotiques pour les fonctions propres, mais seulement des formules pour des fonctions qui sont solutions approchées de l'équation aux fonctions propres. Ces fonctions sont petites en dehors de la projection de la variété lagrangienne sur l'espace de configuration. Les formules asymptotiques présentent des singularités au voisinage des caustiques engendrées par l'application projective.

Les véritables fonctions propres peuvent néanmoins se comporter autrement pour le moins si la valeur propre est multiple ou s'il existe des valeurs propres qui sont proches d'elle (voir Appendice 10).

SINGULARITÉS LAGRANGIENNES

Les singularités lagrangiennes sont les singularités des applications projectives de variétés lagrangiennes sur l'espace de configuration. Ces singularités se rencontrent dans l'étude des solutions de l'équation générale de Hamilton-Jacobi, des caustiques, des points focaux ou conjugués, de la propagation des discontinuités et des ondes de choc en mécanique du milieu continu, de même que dans les problèmes qui conduisent à l'asymptotique à ondes courtes (voir Appendice 11).

Pour décrire les singularités lagrangiennes il faut préalablement dire quelques mots sur les singularités des applications différentiables en général. Commençons par les cas les plus simples.

A. Singularités des applications différentiables d'une surface sur un plan. L'application projective de la sphère sur le plan présente une singularité à l'équateur de la sphère (sur l'équateur le rang de la dérivée baisse d'une unité). On obtient donc sur le plan de projection une courbe (appelée «contour apparent») qui délimite des domaines dont les points n'ont pas le même nombre d'antécédents : chaque point du plan possède deux antécédents à l'intérieur du contour apparent et pas un seul à l'extérieur.

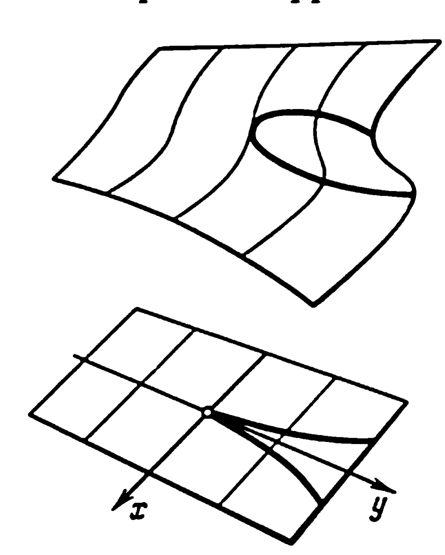


Fig. 245. Singularité de Whitney.

Dans les cas plus généraux le contour apparent est susceptible de présenter des singularités bien plus complexes. Considérons par exemple une surface définie dans l'espace ordinaire muni des coordonnées (x, y, z) par l'équation (fig. 245)

$$x = yz - z^3$$

et l'application projective parallèlement à l'axe z sur le plan muni des coordonnées (x, y).

Les points singuliers de l'application projective forment sur la surface une courbe différentiable (d'équation $3z^2 = y$). Cependant l'image de cette courbe sur le plan (x, y) n'est déjà plus une courbe différentiable. C'est une parabole semi-cubique de sommet au point (0, 0) et d'équation

$$27x^2 = 4y^3.$$

Une telle courbe divise le plan en deux régions: une d'étendue moindre située à l'intérieur et l'autre, plus grande, à l'extérieur. Chaque point intérieur à la courbe est la projection de trois points de la surface, alors que tout point situé à l'extérieur est la projection d'un seul point.

Considérons maintenant une petite déformation quelconque de notre surface. Il s'avère que quelle que soit la surface proche de la nôtre que l'on projette, le contour apparent présentera toujours une singularité analogue (une pointe de parabole semi-cubique) en un certain point proche de la singularité du contour apparent de la surface initiale. En d'autres termes, la *singularité considérée ne disparaît pas par une petite déformation de la surface.*

Bien plus, au lieu de déformer la surface on peut comme bon nous semble déformer l'application projective de la surface sur le plan (quitte à ce qu'elle ne soit plus projective) pourvu qu'elle reste différentiable et que la déformation soit petite. On constate que de telles déformations ne font pas disparaître la pointe, mais la modifient légèrement.

Les exemples cités épuisent les singularités typiques des applications d'une surface sur le plan. On démontre que les singularités plus complexes disparaissent par un petit déplacement. Donc en déformant légèrement une application différentiable quelconque on peut toujours faire en sorte qu'au voisinage d'un point quelconque de la surface projetée elle soit ou bien non singulière, ou bien construite comme une application projective de la sphère sur le plan au voisinage de l'équateur, ou bien comme une application projective de la surface envisagée plus haut avec une pointe semi-cubique sur le contour apparent.

L'expression « construite comme » signifie que sur le plan « but » et la surface « source » on peut choisir des coordonnées locales (au voisinage du point considéré et de son image) telles que l'application s'y écrive sous une forme spéciale. Plus exactement, les formes normales auxquelles est réduite l'application de la surface sur le plan au voisinage des points des trois types indiqués sont

$y_1 = x_1$, $y_2 = x_2$ (point non singulier),

$y_1 = x_1^2$, $y_2 = x_2$ (pli comme sur l'équateur de la sphère),

$y_1 = x_1x_2 - x_1^3$, $y_2 = x_2$ (fronce avec une pointe de contour apparent). Ici (x_1, x_2) sont les coordonnées locales sur la surface « source », (y_1, y_2) celles sur le plan « but ».

La démonstration des théorèmes cités (qui sont dus à H. Whitney) et leurs généralisations figurent dans les travaux portant sur la théorie des singularités des applications différentiables, voir par exemple :

1. V. Arnold — Singularités des applications différentiables (en russe). « Ouspekhi matematitcheskikh naouk », 23, nº 1 (1968), 3-44.

2. R. Thom, G. Levine, J. Mather - Stability of C^∞-mapping, I-IV. Ann. Math., 87 (1968), 89 (1969) 254-291.

B. Singularités de l'application projective des variétés lagrangiennes. Considérons maintenant une variété de configuration de dimension n, l'espace des phases de dimension $2n$ correspondant et une sous-variété lagrangienne de dimension n de cet espace (i.e. une sous-variété de dimension n sur laquelle est identiquement nulle une 2-forme définissant la structure symplectique de l'espace des phases).

En projetant une variété lagrangienne sur l'espace de configuration nous obtenons l'application d'une variété différentiable de dimension n sur une autre de même dimension.

Cette application est un difféomorphisme local dans un point général, cependant le rang de sa dérivée baisse dans certains points de la variété lagrangienne. On dit que ces points sont singuliers. L'ensemble des points singuliers forme, par projection sur l'espace de configuration, un contour apparent que l'on appelle *caustique* dans le cas lagrangien.

Les caustiques sont susceptibles de présenter des singularités complexes, cependant on peut se débarrasser de celles qui sont trop complexes par un petit déplacement comme dans la théorie ordinaire des singularités des applications différentiables. (Par petit déplacement on entend ici une petite déformation qui conserve à la variété sa structure lagrangienne.)

Il ne subsistera que des singularités simples inévitables dont on pourra écrire les formes normales et ensuite en faire une étude détaillée une fois pour toutes. Dans les problèmes génériques qui ne sont doués d'aucune propriété spéciale de symétrie, les seules singularités à se présenter sont naturellement des singularités simples inévitables.

Voyons par exemple les caustiques qui se forment lorsqu'on éclaire un écran avec la lumière d'une source ponctuelle se réfléchissant sur une surface courbe lisse quelconque (ici l'espace des phases de dimension quatre est constitué par des droites coupant la surface de l'écran dans toutes les directions possibles, et la variété lagrangienne, par les rayons issus de la source à leur intersection avec l'écran). En déplaçant la source on constate que d'une façon générale les caustiques ne possèdent que des singularités simples (des pointes semi-cubiques) ; quant aux singularités plus complexes, elles n'apparaissent que pour des positions exceptionnelles de la source.

On cite plus bas les formes normales pour les singularités de l'application projective d'une sous-variété lagrangienne de dimension n d'un espace des phases de dimension $2n$ sur l'espace de configuration de dimension n pour $n \leqslant 5$. Ces formes normales sont en nombre fini et leur classification est (assez mystérieusement d'ailleurs) liée à celles des groupes simples de Lie, des points critiques simples dégénérés des fonctions, des polyèdres réguliers, etc. Pour $n \geqslant 6$ les formes normales de certaines singularités doivent inévitablement contenir des paramètres. Pour plus de détail on renvoie le lecteur à l'article:

V. Arnold — Formes normales de fonctions au voisinage de points critiques dégénérés, groupes de Weyl A_k, D_k, E_k et singularités lagrangiennes (en russe). « Analyse fonctionnelle et ses applications », 6: 4 (1972), 3-25.

C. Table des formes normales des singularités types des applications projectives de variétés lagrangiennes de dimension $n \leqslant 5$. Introduisons les notations suivantes:

$(q_1, \ldots, q_n)$ — coordonnées sur l'espace de configuration,

$(p_1, \ldots, p_n)$ — les impulsions correspondantes de sorte que pris ensemble les p et q forment un système de coordonnées symplectique sur l'espace des phases.

La variété lagrangienne sera définie au moyen de la fonction génératrice F par les formules

$$q_i = \frac{\partial F}{\partial p_i}, \quad p_j = -\frac{\partial F}{\partial q_j},$$

où l'indice i parcourt une partie de l'ensemble $(1, \ldots, n)$ et l'indice j l'autre partie. Plus exactement, $i = 1$, $j > 1$ pour les singularités désignées par A_k et $i = 1, 2$, $j > 2$ pour celles qui sont représentées par D_k et E_k.

Dans ces notations une même expression de la fonction $F(p_i, q_j)$ peut être considérée comme définissant une variété lagrangienne dans des espaces de dimension différente: on peut écrire autant d'arguments q_j que l'on voudra, la fonction F n'en dépendra pas de toute façon.

Voici la liste des formes normales des singularités types:

pour $n = 1$

$$A_1: F = p_1^2; \qquad A_2: F = \pm p_1^3;$$

pour $n = 2$, les deux précédentes plus

$$A_3: F = \pm p_1^4 + q_2 p_1^2;$$

pour $n = 3$, les trois précédentes plus

$$A_4: F = \pm p_1^5 + q_3 p_1^3 + q_2 p_1^2,$$
$$D_4: F = \pm p_1^2 p_2 \pm p_2^3 + q_3 p_2^2;$$

pour $n = 4$, les cinq précédentes plus

$$A_5: F = \pm p_1^6 + q_4 p_1^4 + q_3 p_1^3 + q_2 p_1^2,$$
$$D_5: F = \pm p_1^2 p_2 \pm p_2^4 + q_4 p_2^3 + q_3 p_2^2;$$

pour $n = 5$, les sept précédentes plus

$$A_6: F = \pm p_1^7 \pm q_5 p_1^5 + \ldots + q_2 p_1^2,$$
$$D_6: F = \pm p_1^2 p_2 \pm p_2^5 + q_5 p_2^4 + q_4 p_2^3 + q_3 p_2^2,$$
$$E_6: F = \pm p_1^3 \pm p_2^4 + q_5 p_1 p_2^2 + q_4 p_1 p_2 + q_3 p_2^2.$$

D. Discussion des formes normales. Un point du type A_1 est régulier.

La singularité du type A_2 est une singularité du type pli. En effet, si sur la variété lagrangienne on prend pour coordonnées $(p_1, q_2, \ldots, q_n)$, alors l'application projective s'écrit

$$(p_1, q_2, \ldots, q_n) \mapsto (\pm 3p_1^2, q_2, \ldots, q_n).$$

La singularité du type A_3 est une fronce à pointe semi-cubique sur le contour apparent. Pour s'en assurer il suffit d'expliciter l'application correspondante de la variété lagrangienne de dimension deux sur le plan:

$$(p_1, q_2) \mapsto (\pm 4p_1^3 + 2q_2 p_1, q_2).$$

Les singularités du type A_4 apparaissent pour la première fois dans le cas tridimensionnel et la caustique correspondante est une surface dans l'espace (fig. 246) avec une singularité appelée queue d'aronde (que nous avons déjà rencontrée dans le § 46).

La caustique des singularités du type D_4 est une surface à trois arêtes de rebroussement (du type A_3) tangents en un point; en outre deux de ces trois arêtes peuvent être imaginaires, de sorte qu'il existe deux espaces de caustiques D_4.

E. Equivalences lagrangiennes. Voyons maintenant dans quelle mesure les exemples considérés sont les formes normales des singularités types de l'application projective des variétés lagrangiennes. Définissons au préalable ce qu'on entend par singularités « de même construction ».

L'application projective d'une variété lagrangienne sur l'espace de configuration sera brièvement désignée *application lagrangienne.* Soient données deux applications lagrangiennes de variétés de même dimension n (les variétés lagrangiennes de dimension n correspondantes sont situées en général dans des espaces des phases différents qui sont les fibrés cotangents de deux espaces de configuration différents). Nous dirons que deux telles applications lagrangiennes sont *lagrangiennement équivalentes* s'il existe un difféomorphisme symplectique du premier espace des phases sur le second envoyant une fibre du premier fibré cotangent dans une fibre du second fibré

cotangent et la première variété lagrangienne dans la seconde. Le difféomorphisme en question s'appelle alors *équivalence lagrangienne d'applications.*

Remarquons que deux applications lagrangiennes lagrangiennement équivalentes se transforment l'une dans l'autre par des difféomorphismes dans l'espace source et l'espace but (ou, comme on

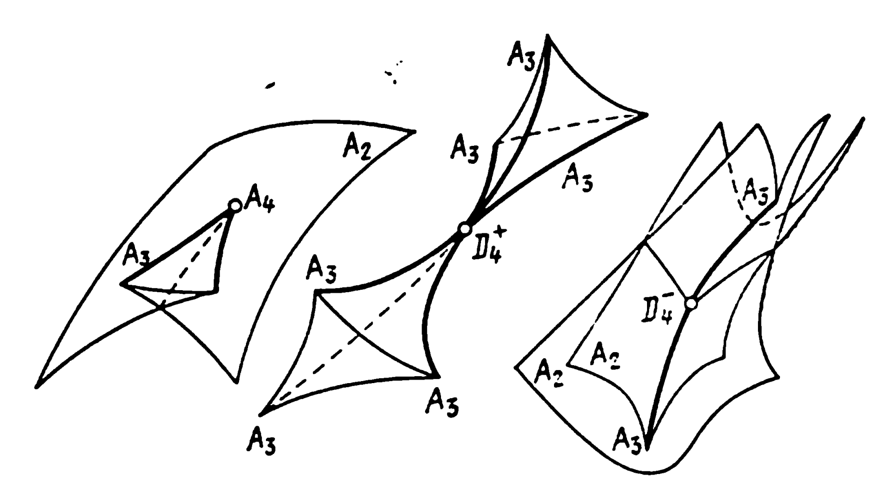


Fig. 246. Singularités types des caustiques dans l'espace ordinaire.

dit en analyse, se transforment l'une dans l'autre par un changement de variables dans l'espace source et l'espace but). En effet, la restriction de notre difféomorphisme à une variété lagrangienne définit un difféomorphisme des sources; le difféomorphisme des espaces de configuration but apparaît parce qu'une fibre se transforme dans une autre.

En particulier, les caustiques de deux applications lagrangiennement équivalentes sont difféomorphes, donc la classification à l'équivalence lagrangienne près entraîne la classification des caustiques. Cependant la classification à l'équivalence lagrangienne près est en général plus fine que la classification des caustiques puisque la difféomorphicité des caustiques n'entraîne pas l'équivalence lagrangienne des applications. Bien plus, la classification à l'équivalence lagrangienne près est plus fine que la classification aux difféomorphismes près de la source et du but, puisque tout tel couple de difféomorphismes n'est pas réalisé par un difféomorphisme symplectique de l'espace des phases.

Une application lagrangienne considérée au voisinage d'un point privilégié est dite *lagrangiennement équivalente en ce point* à une autre application lagrangienne (possédant également un point privilégié) s'il existe une équivalence lagrangienne de la première application dans un voisinage du premier point sur la deuxième application dans un voisinage du second point transformant le premier point dans le second.

Désormais nous pouvons formuler le théorème relatif à la classification des singularités des applications lagrangiennes pour les dimensions $n \leqslant 5$.

Toute variété lagrangienne de dimension n ($n \leqslant 5$) peut être, au moyen d'une déformation aussi petite que l'on veut de la classe des variétés lagrangiennes, transformée dans une autre telle que l'application projective sur l'espace de configuration soit en chaque point lagrangiennement équivalente à l'une des applications lagrangiennes de la liste précédente.

En particulier, une variété lagrangienne de dimension deux peut être, moyennant une déformation aussi petite que l'on veut de la classe des variétés lagrangiennes, transformée dans une variété générique telle que l'application projective sur l'espace de configuration (de dimension deux) ne possède de singularités autres que des plis (qui se ramènent par une équivalence lagrangienne à la forme normale A_2) et des fronces (resp. à la forme normale A_3).

Soulignons que l'assertion relative aux applications lagrangiennes de dimension deux ne découle pas du théorème de classification des applications générales (non lagrangiennes). D'abord parce que les applications lagrangiennes constituent au sein de toutes les applications différentiables une classe très étroite et partant sont susceptibles de posséder (et possèdent réellement pour $n > 2$) en qualité de singularités typiques pour les applications lagrangiennes des singularités qui ne sont pas typiques pour les applications non lagrangiennes. Ensuite le fait que l'on peut réduire une application à la forme normale par des difféomorphismes de la source et du but ne signifie encore pas que cette réduction est possible au moyen de l'équivalence lagrangienne.

Donc les caustiques d'une variété lagrangienne générique de dimension deux admettent comme singularités seulement des pointes semi-cubiques (et des points crunodaux de l'intersection transversale). Les singularités plus complexes se décomposent par une petite déformation de la variété lagrangienne alors que les points de rebroussement et les points crunodaux d'une caustique ne disparaissent pas mais ne font que se déformer légèrement.

Les formes normales des singularités suivantes A_4, D_4, . . . peuvent être de façon analogue appliquées à l'étude des caustiques des variétés lagrangiennes d'un plus grand nombre de dimensions, ainsi qu'à l'étude des modifications des caustiques des variétés lagrangiennes d'un nombre pas très grand de dimensions lorsque varient les paramètres dont dépend la variété.

Les formules du présent Appendice sont utilisées dans la théorie des singularités legendriennes, i.e. des singularités des fronts d'ondes, des transformations de Legendre, des enveloppes et des hulles (voir Appendice 4). La théorie des singularités lagrangiennes et legendriennes trouve des applications immédiates aussi bien en optique géométrique et en théorie des asymptotiques des intégrales oscillantes qu'en calcul des variations, en théorie des solutions discontinues des équations non linéaires à dérivées partielles, dans les problèmes d'optimisation, de poursuite, etc. R. Thom a proposé de grouper la théorie des singularités, des bifurcations et leurs applications sous l'appellation commune de *théorie des catastrophes*.

APPENDICE 13

ÉQUATION DE KORTEWEG-DE VRIES

Les intégrales premières des équations de la mécanique classique ne s'expliquent pas toutes par une symétrie évidente des problèmes (exemple: les intégrales spécifiques du problème de Kepler, du problème des géodésiques sur une ellipsoïde, etc.). Dans ces cas on parle d'une « symétrie sous-jacente ».

L'équation de Korteweg-de Vries

$$u_t = 6uu_x - u_{xxx} \tag{1}$$

fournit d'intéressants exemples de symétrie sous-jacente.

Cette équation non linéaire à dérivées partielles a fait sa première apparition en théorie de l'étiage; par la suite on l'a retrouvée dans une foule de problèmes de physique mathématique.

De nombreuses expériences numériques ont mis à jour les étonnantes propriétés des solutions de cette équation satisfaisant à des conditions aux limites nulles à l'infini : lorsque $t \to +\infty$ et $t \to -\infty$, ces solutions se décomposent en « solitons », i.e. en ondes de forme définie progressant à des vitesses différentes.

Pour obtenir un soliton progressant à la vitesse c il suffit de porter la fonction $u = \varphi(x - ct)$ dans l'équation (1). L'on a alors l'équation $\varphi'' = 3\varphi^2 + c\varphi + d$ (d est un paramètre). C'est une équation de Newton à potentiel cubique. Il existe un point selle sur le plan de phase $\{(\varphi, \varphi')\}$. La séparatrice qui joint un point selle à celui où $\varphi = 0$ définit une solution φ qui tend vers 0 lorsque $x \to \pm\infty$; cette solution est un soliton.

Lorsque les solitons se heurtent, on observe une interaction non linéaire relativement complexe. Cependant une expérience numérique a montré que les dimensions et vitesses des solutions ne changeaient pas après une collision. Ce qui a suggéré l'idée de lois de conservation. Et effectivement, M. Kruskal, Zabousski, P. Lax, C. Gardner, J. Green et R. Miura ont réussi à trouver toute une série d'intégrales premières pour l'équation de Korteweg-de Vries. Ces intégrales sont de la forme $I_s = \int P_s(u, \ldots, u^{(s)})\, dx$, où P_s est un polynôme. Il est aisé de vérifier que les intégrales premières

de l'équation (1) sont

$$I_{-1} = \int u\,dx, \quad I_0 = \int u^2\,dx, \quad I_1 = \int \left(\frac{u'^2}{2} + u^3\right) dx,$$

$$I_2 = \int \left(\frac{u''^2}{2} - \frac{5}{2}u^2u'' + \frac{5}{2}u^4\right) dx.$$

L'apparition d'une série infinie d'intégrales premières s'explique aisément par le théorème suivant de Lax *). Désignons l'opérateur de multiplication d'une fonction de x par le symbole de cette fonction, et l'opérateur de dérivation par rapport à x par ∂. Considérons l'opérateur de Sturm-Liouville $L = -\partial^2 + u$ qui dépend de la fonction $u(x)$. Il est immédiat de vérifier le

Théorème. *L'équation de Korteweg-de Vries* (1) *est équivalente à l'équation* $\dot{u} = [L, A]$, *où* $A = 4\partial^3 - 3(u\partial + \partial u)$.

Du théorème de Lax il découle aussitôt le

Corollaire. *Les opérateurs L construits d'après une solution de l'équation* (1) *sont unitairement équivalents pour tous les t; en particulier, chacune des valeurs propres λ du problème de Sturm-Liouville $Lf = \lambda f$ avec des conditions nulles à l'infini est intégrale première de l'équation de Korteweg-de Vries.*

V. Zakharov et L. Faddéev ont remarqué que l'équation (1) est un système hamiltonien de dimension infinie entièrement intégrable et ont indiqué les variables action-angle **) correspondantes. Une structure symplectique sur l'espace des fonctions $u(x)$ décroissantes à l'infini est donnée par le produit scalaire gauche $\omega^2(\partial w, \partial v) = \frac{1}{2}\int (w\,dv - v\,dw)\,dx$; quant au hamiltonien de l'équation (1), c'est l'intégrale I_1. En d'autres termes, l'équation (1) prend la forme de l'équation de Hamilton $\dot{u} = \frac{d}{dx}\frac{\delta I_1}{\delta u}$ dans l'espace des fonctions de x.

Toute intégrale I_s définit donc une « équation supérieure de Korteweg-de Vries » $\dot{u} = Q_s[u]$, où $Q_s = \frac{d}{dx}\frac{\delta I_s}{\delta u}$ est un polynôme de $u, u', \ldots, u^{2s+1}$. Les intégrales I_s sont en involution et les flots qui leur correspondent dans l'espace fonctionnel commutent.

Les polynômes P_s et Q_s, ainsi que les variables action-angle (et partant les solutions de l'équation (1)) s'explicitent en termes de solutions du problème direct et du problème inverse de la théorie de dissipation sur le potentiel u.

*) P. Lax, *Intégrales des équations d'évolution et ondes isolées* (en russe). Recueil « Mathématiques », 13: 5 (1969), 128-150.

**) V. Zakharov, L. Faddéev, *L'équation de Korteweg-de Vries est un système hamiltonien entièrement intégrable* (en russe). « Analyse fonctionnelle», 5: 4 (1971), 18-27.

On peut expliciter les polynômes Q_s en utilisant le théorème suivant de Gardner qui est une extension de celui de Lax. Soit dans l'espace des fonctions de x un opérateur différentiel de la forme $A = \sum p_i \partial^{m-i}$, où $p_0 = 1$, les autres coefficients p_i étant des polynômes de u et u'_x. Il s'avère que *pour tout s il existe un opérateur A_s d'ordre $2s + 1$ tel que son commutateur avec l'opérateur L de Sturm-Liouville soit l'opérateur de multiplication par une fonction*: $[L, A_s] = Q_s$.

L'opérateur A_s est défini par les conditions ci-dessus de façon unique à l'addition près d'une combinaison linéaire A_r, où $r < s$; donc les polynômes Q_s en u et u'_x sont définis à l'addition près d'une combinaison linéaire des Q_r.

V. Zakharov, A. Chabat, L. Faddéev et d'autres ont utilisé le procédé de Lax et la technique du problème inverse de la théorie de la dissipation pour étudier un grand nombre d'importantes équations de la physique et notamment les équations $u_{tt} - u_{xx} = \sin u$, $i\psi_t + \psi_{xx} \pm \psi \mid \psi \mid^2 = 0$.

L'étude du problème à conditions limites périodiques pour l'équation de Korteweg-de Vries a permis à S. Novikov *) de découvrir une intéressante classe de systèmes complètement intégrables à nombre fini de degrés de liberté. Ces systèmes se construisent de la manière suivante.

Considérons une combinaison linéaire finie quelconque d'intégrales premières $I = \sum c_i I_{n-i}$ et soit $c_0 = 1$. L'ensemble des points stationnaires du flot de hamiltonien I dans l'espace fonctionnel est invariant par les flots de hamiltoniens I_s, notamment par le flot de l'équation (1).

D'un autre côté ces points stationnaires se définissent à partir de l'équation $\frac{d}{dx}\frac{\delta I}{\delta u} = 0$ ou $\frac{\delta I}{\delta u} = d$. On reconnaît dans la dernière équation l'équation d'Euler-Lagrange pour la fonctionnelle $I - dI_{-1}$ qui comprend les dérivées n-èmes. Elle est donc d'ordre $2n$ et peut être écrite comme un système d'équations de Hamilton dans un espace euclidien de dimension $2n$.

Il s'avère que le système hamiltonien à n degrés de liberté obtenu possède n intégrales en involution et peut être totalement intégré moyennant un choix des coordonnées action-angle. On obtient donc une famille de dimension finie de solutions particulières de l'équation de Korteweg-de Vries dépendant de $3n + 1$ paramètres ($2n$ coordonnées et $n + 1$ paramètres $c_1, \ldots, c_n$; d).

Novikov a montré que les solutions trouvées étaient douées de remarquables propriétés: dans le problème périodique par exemple elles définissent des fonctions $u(x)$ pour lesquelles l'équation différentielle à coefficients périodiques $-X'' + u(x)X = \lambda X$ possède un nombre fini de zones de résonance paramétrique (voir § 25) sur l'axe λ.

---

*) S. Novikov, *Problème périodique pour l'équation de Korteweg-de Vries*, I (en russe). « Analyse fonctionnelle », 8: 3 (1974), 54-66.

INDEX ALPHABÉTIQUE DES MATIÈRES

